

In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institution shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

10

A COMPUTATIONAL TECHNIQUE FOR IDENTIFICATION
OF LINEAR SYSTEMS

A THESIS

Presented to
The Faculty of the Graduate Division
by
Clifford Overall Guffee, Jr.

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the School
of Electrical Engineering

Georgia Institute of Technology
August, 1965

A COMPUTATIONAL TECHNIQUE FOR IDENTIFICATION
OF LINEAR SYSTEMS

Approved: _____

Dr. J. L. Hammond, Jr., Chairman

Dr. R. P. Webb

Dr. D. L. Finn

Date approved by Chairman September 3, 1965

ACKNOWLEDGMENTS

Part of the satisfaction that comes of doing research is derived from association with one's associates, and it is natural that through the exchange of views and the "talking over" of problems that one's work reflects the ideas of many people. A list of such persons for this research would be long, indeed, and credit can be given to only a few. First of all, the writer would like to thank Dr. J. L. Hammond, Jr., under whose supervision this research was carried out, for his patience, and encouragement throughout the preparation of this manuscript, and Dr. R. P. Webb and Dr. D. L. Finn for their interest and practical suggestions. He would like also to acknowledge the Rich Electronic Computer Center first for providing the computer time so necessary for the full development of this research, and second for the aid and assistance of its staff during his programming of the problem.

He would also like to express special appreciation to the National Aeronautical and Space Administration and to the National Science Foundation for the separate fellowships during the period while this research was conducted.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.	ii
LIST OF TABLES	v
LIST OF ILLUSTRATIONS.	vi
SUMMARY.	vii
Chapter	
I. INTRODUCTION	1
Available Identification Procedures	
Uncertainty in the Identification Problem	
Outline of this Study	
II. FORMULATION OF THE IDENTIFICATION PROCEDURE	10
Identification of a Model	
Identification Performance Index	
Specification of the Identification Problem	
Introduction to the Identification Procedure	
Development of a System of Linear Equations	
Defining the Identification Procedure	
Formation of the Normal Equations	
Determination of Order	
Response Error and the Least Squares Solution	
Class of Allowable Excitation	
III. ERROR ANALYSIS	35
Source of Errors	
Inherent Errors	
Errors in Numerical Integration Methods	
Roundoff Errors and Ill-Conditioned Matrices	
Uncertainty in the Parameters	
IV. GENERAL CONSIDERATIONS IN THE IMPLEMENTATION OF THE IDENTIFICATION PROCEDURE	50
Introduction	
Obtaining Test Data	
Errors Introduced by Numerical Integration	

Chapter		Page
IV.	GENERAL CONSIDERATIONS IN THE IMPLEMENTATION OF THE IDENTIFICATION PROCEDURE (continued)	
	Solving a System of Simultaneous Equations	
	Uncertainty in the Results of the Identification Procedure	
	Determination of Order	
	Surplus Factors	
	The Condition of Different Systems of Equations	
V.	IDENTIFICATION: GIVEN IMPULSE RESPONSE.	80
	Response Length Required for Identification	
	Establishing the System of Equations	
	Gradient of the Error Function	
	Automation of the Identification Procedure	
	Examples	
	Conclusions	
VI.	MODEL OF NEUTRON DECAY WITHIN A REACTOR.	115
	Problem Statement	
	Solution Using the Identification Procedure	
	Solution Using a Multivariable Search Routine	
	Conclusions	
VII.	IDENTIFICATION FOR A GENERAL EXCITATION.	125
	Introduction	
	Automating the Identification Procedure	
	Outline of the Investigation	
	Results of the Identification Attempts	
	Conclusions and Recommendations	
VIII.	CONCLUSIONS.	137
Appendices		
A.	MATHEMATICAL MODEL	143
B.	IDENTIFYING TIME-VARYING SYSTEMS	153
C.	NORMED ERROR ANALYSIS.	163
D.	SOLVING A SYSTEM OF LINEAR ALGEBRAIC EQUATIONS	171
E.	IDENTIFICATION USING MULTIVARIABLE SEARCH.	179
BIBLIOGRAPHY		185
VITA		189

LIST OF TABLES

Table	Page
1. Characteristics of a Test System.	54
2. Comparison of Integrating With Different Step Intervals .	60
3. Uncertainty in the Solution of a System of Equations. . .	70
4. Comparison of Results Using Two Different Sampling Intervals	71
5. Comparison of Results of Identification Attempts Using Essentially True Data	89
6. Results of Iterating on a Solution After Surplus Factors Have Been Removed	94
7. Comparison of Results for Unsuccessful Attempts	95
8. Effects of Additive Noise for $B = 0.0025$	100
9. Comparison of Uncontaminated Signal and Additive Noise for $B = 0.0025$	101
10. Comparison of Results for $B = 0.0025$	102
11. Effects of Additive Noise for $B = 0.005$	103
12. Comparison of Uncontaminated Signal and Additive Noise for $B = 0.005$	103
13. Comparison of Results for $B = 0.005$	104
14. Effects of Additive Noise for $B = 0.025$	104
15. Comparison of Uncontaminated Signal and Additive Noise for $B = 0.025$	106
16. Comparison of Results for $B = 0.025$	107
17. Characteristics of a Third Order System	128
18. Frequency Spectrum of Test Excitations.	131

LIST OF ILLUSTRATIONS

Figure	Page
1. Uncertainty in the Identification Problem.	7
2. Sources of Error.	36
3. Plot of Impulse Response of Test System.	55
4. Block Diagram of Basic Identification Procedure and Input Simulator	85
5. Plot of Neutron Decay	116
6. Frequency Plot of a Third Order System.	129
7. Frequency Plot of a Fourth Order System	130

SUMMARY

The purpose of this study is to develop a computational technique for the identification of a class of linear systems. The identification procedure will determine the parameters of a differential equation which describes an unknown system based only upon information derived from measurements made on the excitation and response of the system. The only a priori knowledge about the system is that it is linear, time-invariant, stable, and has either a low-pass or a band-pass frequency characteristic. Specifically, no knowledge need be available as to the order of the system, the value of the initial conditions within the system, or the value of any coefficients in the differential equation which describes the system.

The identification procedure is developed mainly for the case where the system is excited by an unit impulse, but it is also applicable for the problem of identifying a system for the case where the excitation is a member of a larger class of signals. The theoretical derivation of the identification procedure is for the case where the system is excited by a general signal, and after the procedure has been fully developed from a theoretical standpoint, the aspects of the identification problem concerned with the two types of excitation are considered as separate problems. In particular, consideration is given to such problems as the minimum required time length of available response and the minimum accuracy of the measured data which will still allow the identification procedure to correctly identify a system.

Consideration is also given to the requirements upon the frequency spectrum of an excitation (for the case where the excitation is other than an unit impulse), which will allow successful identification.

The identification procedure is developed theoretically for the situation of exact data and exact computational methods. It is formulated as a problem of solving a system of linear algebraic equations where the system of equations ~~is~~ established using data obtained only from measurements made on the excitation and response of the system. In particular, the required data includes measurements of both the excitation (for the case where the excitation is other than an unit impulse), the corresponding response, and integrated values of these signals. The required integration is an n^{th} order process where n is at least equal to the order of the system being identified.

In the theoretical development of the identification procedure it is proved that the solution of the identification problem will be an exact model of the unknown system. It is also shown that no a priori knowledge of the properties of the system is required in solving the identification problem. An error analysis of the theoretical identification procedure considers the problem which exists in practice because of the necessity of working with inexact data and inexact methods. In this analysis it is shown that the error between the parameters of a system and the parameters of a determined model is bounded by an expression involving the inaccuracy of the available data and the errors introduced by the required numerical methods. Consideration is given to the problem of minimizing the inaccuracies introduced by the numerical methods, and procedures are presented for

estimating a bound on the uncertainty of a solution of a practical identification problem. The bound on the uncertainty is determined from knowledge of the accuracy of the data as well as from knowledge of the errors introduced by the required numerical methods.

The basic identification procedure is implemented on a digital computer for the case where the excitation is an unit impulse, in order to, first, develop a working procedure and, second, to investigate several aspects of the procedure. The implementation considers the problem of how to best establish and solve the required system of equations as well as the physical problem of determining the uncertainty of the final solution. The investigation section considers the problem of the minimum required time length of the response required for successful identification as well as the problems encountered when the data is contaminated by noise. In particular, the investigation considers the problem of determining the correct order of the system when solving a practical problem. Examples are presented to illustrate all significant aspects of both the implementation and the investigation, and all results are related to theory.

As an example of identification attempts with accurate data, an example is given in which the impulse response of a fourth order system is sampled and then integrated by the trapezoidal rule. The sampled data has approximately five significant digits and the integrated data has approximately three significant digits. All identification attempts using this data resulted in a correct order model, and in general the individual coefficients of the models are within one per cent of the respective coefficients of the true system.

The above samples are disturbed by a random factor between ± 2.5 per cent of the sample value being contaminated. The contaminated data has one to two significant digits at each sample value. The samples are again integrated by the trapezoidal rule and identification is attempted using the data. Even with such contaminated data, the identification procedure still determined the correct order of the model although the parameter values are not determined with the same accuracy as the previous example.

In the ideal case the identification procedure will minimize the RMS error between the response of the system and the response of the determined model which is to represent the system. However, because of the inexactness in the data, the practical identification procedure does not truly minimize the RMS error. For relatively accurate data the identification procedure can yield an essentially true model, and with relatively inaccurate data the identification procedure can still yield a good estimate for the parameters of the correct order model. Thus, it is possible to use the identification procedure to determine the correct order of a model and good initial estimates of its parameter values. This model can then be refined with a multivariable search procedure. The extent of refinement that is necessary will depend upon the extent of the inaccuracy of the data. Discussions and examples are presented to illustrate significant aspects of the combined use of the two identification procedures as well as to illustrate the results which can be obtained when using data with various degrees of accuracy.

The identification procedure is compared to an identification method which employs a multivariable search routine. It is shown that

the identification procedure and multivariable search routines complement each other. A multivariable search routine requires a priori knowledge as to the order of the model as well as initial estimates of the values of the parameters of the model. The parameters are then adjusted so as to minimize the RMS error between the response of the model and the system which it represents. On the other hand, the identification procedure of this research requires no a priori knowledge as to the order or as to the value of the parameters of the model.

Data expressing the neutron flux decay within a nuclear reactor, after the reactor has been shut down by a reactive step, has been experimentally obtained from the nuclear reactor at the Georgia Institute of Technology. The identification procedure has been applied to the practical problem of determining a differential equation model which describes this initial condition response of the reactor. The steps involved in solving this practical identification problem are given, and results are compared to results obtained during the investigation of the identification procedure.

A separate implementation and investigation of the identification procedure is conducted for the case where the system excitation is a general signal. In this investigation it is determined that there must be restrictions upon the frequency spectrum of the excitation signal; however, the other aspects of the identification procedure are still general -- that is, the identification procedure can still determine the order of a model and relative values of its coefficients. As before, the accuracy of the determined coefficients depends upon the accuracy of the data. Examples and discussions are included on this aspect of the identification procedure.

CHAPTER I

INTRODUCTION

Identification can be described briefly as the gathering and combining of information about a specific process into a form which will allow the user of the model to predict the future responses of the process. There is, of course, a common denominator for all attempts at identifying a process. Some properties of the excitations and corresponding responses of the process must be combined together in some fashion to yield approximate, numerical information regarding the dynamics of the process. A completely general identification procedure would not require prior information on properties of the process; however, in some situations there can be available information, such as process order, the form of the model, approximate value of parameters, and existing initial conditions, which can allow the use of an identification procedure that requires such a priori information.

The solution of the identification problem is particularly important in the design of feedback control systems. In control system design, the dynamic characteristics of a process to be controlled must be known in order to effectively carry out the design of the controller. If the process exhibits parameter variation, the solution of the identification problem takes on a different significance. Processes which exhibit parameter variation have placed increased emphasis on the concept of adaptive control--or control in which automatic and continual measurements of the dynamics of the process to be controlled is used as a basis

for the automatic and continuing self-design of the controller. Adaptivity then implies automatic, frequent, and rapid solution of the identification problem.

Available Identification Procedures

The identification problem as related to feedback control systems can be solved by several processes. The most direct is by use of physical laws governing dynamic behavior. As an example, consider the case of a system using a d-c motor, and driving through a gear train, a load consisting of damping and inertia. This system can in general be described by a set of six differential equations which can be formulated from theoretical laws.* In other situations, the physical laws will not be known and other procedures and models are required.

The impulse response of a process is a complete description which can be found in several ways. One way is by use of crosscorrelation, as first introduced by Lee.** Lee pointed out that when the input of a linear, time-invariant system is white noise, the crosscorrelation function between the input and output is the system's weighting function. This method has been applied to the identification problem of adaptive systems by Anderson, Buland, and Cooper⁽¹⁰⁾.

Identification techniques are also available which estimate the impulse response of an unknown linear, time-invariant system from data obtained by sampling the system's input and output signals. Levin⁽¹¹⁾

* Mishkin and Braun (5), pp. 59-66.

** Lee (8), also Lee (9), pp. 342-348.

considered this method using statistical estimation theory when the output is corrupted by noise. This same approach is used also by Kerr and Surber⁽¹²⁾. In both cases, the result of the identification procedure is a specification of the impulse response of the system. Other authors considering identification based upon determination of the impulse response are Hsich⁽⁷⁾, who gives a least squares estimation which is applicable for both linear and nonlinear systems, and Turin⁽¹³⁾ who estimates the impulse response in the presence of noise by a match filter technique.

While the determination of the impulse response of a system does identify a system, in some cases it is not as useful as other models such as a differential equation. The identification of a process in the sense of determining a vector differential equation has been considered by such authors as Breman, Shahn, and Weiss⁽¹⁵⁾. These authors studied the formation of models to fit certain empirical data for the study of biological systems. They considered models of linear, time-invariant systems, and their paper deals chiefly with calculating the values of the parameters of the model. Their method is not entirely general in that the order of the system, the initial conditions and initial estimates of the value of the parameters must be given. Another paper⁽¹⁴⁾ by the same authors discusses the stages of model building when the model is to be a vector differential equation.

Another approach to the identification problem is given by Bellman, Kagiwade and Kalaba⁽¹⁶⁾. These authors assume that the general form of the dynamic equation is known, although the exact equation is unknown. Observations are made on the response of the system, and initial estimates are made as to the value of the parameters. An iteration method

is employed to obtain the actual value of the parameters. An example is given for a nonlinear system in which different initial approximations to the value of the parameters were tried. In most cases, the procedure yielded very satisfactory results.

Two other authors, Kumar and Sridhar⁽¹⁷⁾, also have a procedure for identifying the coefficients of the differential equation describing a linear, time-invariant process. Their procedure considers the coefficients as random variables and through the use of statistical estimation theory, they are able to determine the coefficients if the form, and order of the differential equation is known.

Surber⁽²⁴⁾ discusses identification based upon a differential equation model in which the coefficients are estimated from the observed set of input and output variables. Two techniques are outlined, and in both cases the model structure, including the order of the system is assumed to be known. A multivariable search procedure is then employed to determine the parameters of the model. This procedure is applicable for any process for which a model and initial estimates of the value of the parameters of the system are known.

Diamesis^{(45),(46)} has commented upon a identification procedure which is similar to the work of this research and which is applicable for any process for which a model is available. His articles present only the theoretical procedure, and it is not as yet a working method. A further comparison of Diamesis' work and the work of this research is given in Chapter II.

Uncertainty in the Identification Problem

A theoretical identification procedure is based on the premise that exact data is available and that exact methods can be employed to implement the procedure. Here, data includes all information which the identification procedure requires about the unknown system, while methods include any computational process which the identification procedure must use in obtaining a solution. This theoretical identification procedure will then have exact rules which will always lead to an exact solution in theory.

Throughout this work, it is necessary to use adjectives such as exact, ideal, etc., and they will be used to describe a situation which can occur only in theory. Similarly, adjectives such as inexact, non-ideal, etc., will be used to describe situations which occur in physical practice. In solving an identification problem with a computational identification procedure it is necessary to work with inexact data and inexact methods. In the actual case, the available data (and methods) can vary from one extreme of being physically meaningless (or useless) to the other extreme of being arbitrarily close to exactness.

An uncertainty in the obtained solution of an identification problem will occur because of inexactness in first, the required methods and second, the available data. A simple example of the uncertainty caused by each of these sources of error occurs when it is necessary to use the value of Π in a computation. If the value of Π is available to any desired number of digits but the required methods will allow only ten digits to be used, then the inexactness in the solution is due primarily to the methods. On the other hand, if only a three digit

approximation to the value of Π is available, then the data is the major factor limiting the accuracy of the solution. While this is a rather elementary example, it illustrates the type of problems encountered in all computational work. It is necessary in practice to be able to estimate the uncertainty in a solution from a knowledge of the accuracy of the data, and the inexactness of the methods employed to obtain this solution.

Figure 1 is a hypothetical plot of "Uncertainty in Solving the Identification Problem" versus "Accuracy of the Data." The abscissa represents the accuracy of the required data, and its range is from 0--or no data--to ∞ -- or exact data. Thus, the available data in practice will lie somewhere between these two extremes. The ordinate represents total uncertainty which includes both uncertainty as to the applicability of the identification procedure, as well as the uncertainty as to the reliability of an obtained solution. This coordinate is thus scaled from 0--or the exact situation with no uncertainty--to ∞ -- or complete vagueness. At the upper limit, the identification procedure is no longer applicable, while at the lower limit the procedure is considered only in a theoretical sense.

Somewhere on the abscissa of this hypothetical plot there is a range of minimum accuracy in the data for which the identification procedure is still applicable. For available data within this range, a solution is obtained not by applying specific rules, but by using intuition and a working knowledge of the required methods. As the accuracy of the data increases, the uncertainty of the problem decreases -- that is, fixed rules can be given as to how the identification

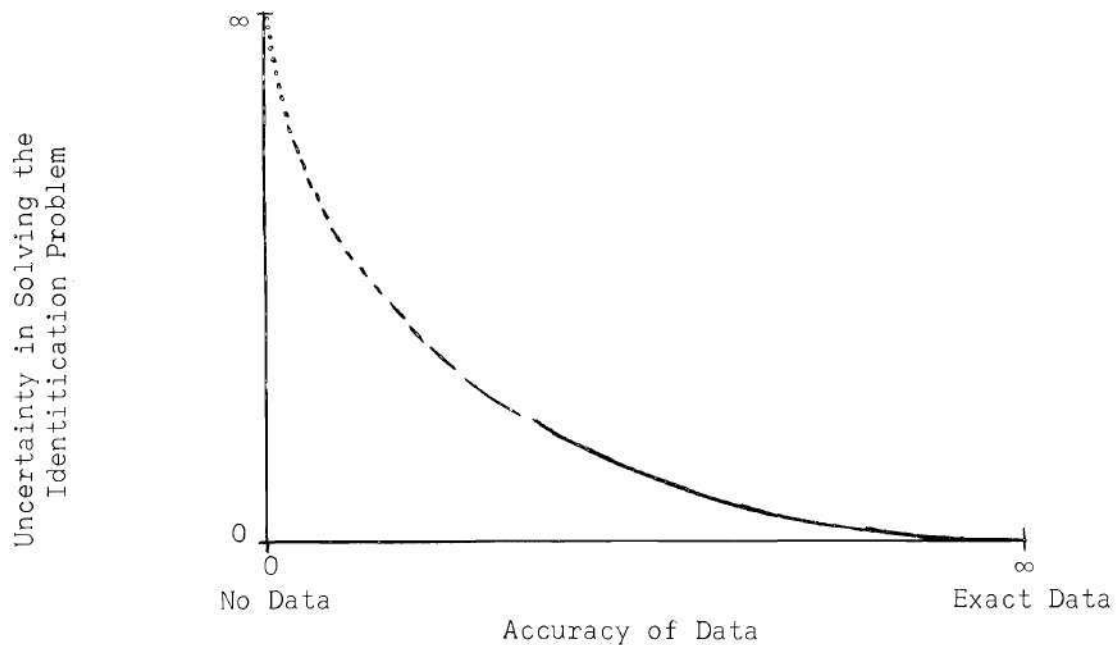


Figure 1. Uncertainty in the Identification Problem.

procedure is applied. As the accuracy of the data becomes more exact, the confidence which can be placed in the solution is limited only by the inexactness of the required methods. This does not mean that a good solution can only be obtained with near-exact data and methods, but rather that the user must realize the limitations of the procedure.

Thus while it is necessary that a computational identification procedure provide step-by-step instructions for the solution of the identification problem and for estimating the uncertainty in the obtained solution, it is also necessary to provide the user with an insight into the limitations and flexibilities of the methods used within the procedure. In this way, the user can then adapt the procedure to fit his needs and/or limitations both with respect to the accuracy of the available data, and with respect to the facilities with which the procedure is to be implemented.

Outline of this Study

The problem considered in this study is that of developing a computational technique for identifying a class of linear systems by determining a differential equation for the system when given the impulse response of the unknown system. The only a priori knowledge about the system is that it is stable, linear, time-invariant and has either a low-pass or band-pass frequency characteristic. Specifically, no knowledge need to be available as to the order of the system, the value of initial conditions, or the value of the coefficients in the differential equation which describes the system. The data required by this identification procedure can be obtained from measurements made upon the response of the system.

The development of the identification procedure is extended over the next four chapters, each chapter being devoted to a particular aspect of the problem. These chapters can be summarized as follows:

1. Chapter II contains the theoretical development of the identification procedure, and indicates the general problems which will exist when this procedure is used in practice. Further discussions are also included on the general problem of identification.

2. Chapter III contains an error analysis of the computational methods used within the identification procedure. In particular, the analysis considers the problem of determining the uncertainty of an obtained solution based upon a knowledge of the uncertainty of the data, and the inaccuracies of the numerical methods employed within the procedure. Again, discussions are included on the problems which will be encountered in practice.

3. Chapters IV and V contain both an implementation and an investigation of the identification procedure for the case where the system is excited by an unit impulse. The procedure is automated, and conclusions are formed based upon actual identification attempts using test data. Examples are included in order to illustrate significant results, and all results are related to theory. Chapter V contains a summary of the complete identification procedure.

4. The developed identification procedure has been applied to a practical problem. The flux decay within a nuclear reactor has been measured, and the identification procedure is employed to determine a model for this decay. Chapter VI contains this discussion, and all aspects of the attempted identification are related to the discussions and theory of the previous chapters.

5. While the identification procedure is considered mainly for the case where the system has been excited by an unit impulse, it is also applicable for the problem of identifying a system where the excitation includes a larger class of signals. The theoretical derivation of the identification procedure in Chapters II and III is thus for the case where the system has been excited by a general signal $y(t)$. After the identification procedure has been examined for the special case of $y(t)$ being an impulse, consideration is then given to the more general aspect of the identification problem. Chapter VII then contains an investigation of the identification procedure when $y(t)$ is other than an impulse.

CHAPTER II

FORMULATION OF THE IDENTIFICATION PROCEDURE

This chapter contains the theoretical development of an identification procedure. A general scalar differential equation representation of a class of systems is analyzed for the situation where a system is excited by a general signal $y(t)$, and the identification problem is reduced to that of establishing and solving a system of linear algebraic equations. The only information required by the identification procedure can be obtained from measurements made on the excitation and response of the system.

Identification of a Model

While the intent of identification is to determine the parameters of an unknown system, the actual results of an identification procedure will be the specification of the parameters of a model. The object of any identification process is then to develop a model which is as close as possible to a true representation of the unknown system. If the measurements of the excitation and response of a system are exact, and if the identification procedure uses exact methods, the resulting model will then be exact. With inexact data and approximate methods, the resulting model will then be only an estimate of the true system. In this case, the identification procedure must also specify the uncertainty of the parameters of the determined model.

When the required measured data is subject to error, such as will be the situation when the response of the system is contaminated by noise,

it is then necessary to match the response of the model with the measured response of the system at a large number of sample times. Further, these samples must be taken over a long period of time so as to minimize the effect of noise. If the chosen samples are representative, then the response of a model, which is determined by the procedure, will approximate the response of the system not only at the sample times, but also over the total response time outside the interval over which the samples have been taken.

Identification Performance Index

Identification as discussed in this research is then the determination of the parameters of a model so that the response of this model is as close as possible to the measured response of the unknown system when the system and the model have the same excitation. In this section, a method of judging closeness is given, and this will be the criterion used throughout this research whenever discussing the closeness of a model with respect to the system it is to represent.

Assume that a system of order n has a measured response given by $x(t)$, where the measurement is made at the indicated time. If a model of order s has a response of $z(t)$ when excited by the same signal that excites the system, it is reasonable to define an instantaneous error by

$$E(t;y) = x(t) - z(t), \quad (2.1)$$

where the error is a function of both time and the particular excitation $y(t)$.

If the responses of both the system and the model are available as continuous signals, the root-mean-squared error is then defined as

$$\text{ERMS} = \left\{ \frac{1}{T} \int_0^T |E(t;y)|^2 dt \right\}^{1/2}, \quad (2.2)$$

where the interval $[0,T]$ represents the interval for which the responses can be measured. A performance index, I , will be defined as the normalized error, and the normalizing factor must be chosen so as to indicate the relative size of the ERMS error with respect to the size of the response of the system over the same interval. The performance index will be defined as

$$I = \frac{\text{ERMS}}{\left\{ \frac{1}{T} \int_0^T |x(t)|^2 dt \right\}^{1/2}}. \quad (2.3)$$

For the case where the response of the system has been measured only at T discrete times, the mean-squared factors of Equations 2.5 and 2.6 will be taken as the average of the sum of squares of the instantaneous errors. Thus Equation 2.2 becomes

$$\text{ERMS} = \left\{ \frac{1}{T} \sum_{i=1}^T |E(t_i;y)|^2 \right\}^{1/2}, \quad (2.4)$$

where T represents the number of discrete times at which the response of the system has been measured. Similarly the performance index becomes

$$I = \frac{\text{ERMS}}{\left(\frac{1}{T} \sum_{i=1}^T |x(t_i)|^2 \right)^{1/2}} \quad (2.5)$$

It must be understood that the ERMS error criterion is a measure of the closeness of fit between the response of the model and the response of the system for fixed excitation. This criterion does not measure the deviation between the parameter values in the equations describing the system and the model. Hopefully, if two models are both possible approximate representations of the same system, then the model which exhibits the smallest ERMS error will then also be described by an equation whose parameter values are closer to those of the equation of the actual system. However, examples will be given in later chapters of cases for which one possible model will have a smaller ERMS error than a second model, but the second model has parameter values which are better approximations of the parameters of the actual system.

An identification procedure can employ multivariable search routines and be directly concerned with minimizing ERMS; however, the identification procedure of this research does not directly minimize this error function, although this is the desired result. This point will be explained in a later section.

Specification of the Identification Problem

A class of stable, linear, time-invariant, n^{th} order systems can be described by a linear, constant coefficient differential equation of the form

$$L_n x(t) = N_n y(t), \quad (2.6)$$

along with initial conditions $x(0), \dot{x}(0), \dots, \dot{x}^{(n-1)}(0)$. In Equation 2.6, $x(t)$ is the system output variable, $y(t)$ is the excitation variable, and the operators L_n and N_n are defined by

$$L_n = \sum_{i=0}^n p_i \frac{d^i}{dt^i}, \text{ with } p_n = 1, \quad (2.7)$$

$$N_n = \sum_{i=0}^{n-1} q_i \frac{d^i}{dt^i}. \quad (2.8)$$

The object of this research is to develop a computational technique for determining the differential equation of an unknown system which belongs to the class of systems which can be described by Equation 2.6. This restricts the class of systems to those systems which are linear, time-invariant, and which have either a low-pass or band-pass frequency characteristic. Equation 2.6 will be referred to as describing an n^{th} order system, and the equation has been normalized so that the leading coefficient, defined by p_n , is unity. The remainder of the coefficients are restricted only by the requirement that they be finite in value. Appendix A contains a discussion on a mathematical model in the form of a vector rather than a scalar differential equation, and the material of that appendix will be used in the derivation of the identification procedure.

The identification procedure will determine the order of the required differential equation, as well as the value of its coefficients

and initial conditions from data obtained by measuring the response and excitation of the system. A theoretical development of the identification procedure is given for the case of a general excitation $y(t)$, and a practical computational procedure is developed in detail. The theoretical development of this chapter as well as the discussion of Chapters III and IV will be for the case of a general excitation. Chapter V contains the final development of the procedure where the excitation is a unit impulse, and Chapter VII will discuss identification for excitations other than an impulse. In particular, the necessary restrictions upon a general excitation are given in Chapter VII.

The identification procedure is also applicable for the problem of identifying the parameters of systems which have time-varying coefficients. For this situation it is necessary that several restrictions be satisfied, and Appendix B discusses this aspect of the identification problem.

Introduction to the Identification Procedure

Since Equation 2.6 represents a general mathematical model of the class of systems considered in this research, it is to be investigated with the idea of establishing a system of linear algebraic equations such that the solution of this system of equations will yield the order, the values of the coefficients and the values of the initial conditions of the model. Further, the simultaneous equations are to be established with information which can be determined from measurements made on the excitation and response of the unknown system.

An obvious method for establishing the desired system of equations is to measure the excitation and response and the derivatives of these signals at enough points to establish the required system of equations. The j^{th} equation will then have the form

$$\frac{d^n x(t_j)}{dt^n} = \sum_{i=0}^{n-1} \hat{p}_i \frac{d^i x(t_j)}{dt^i} + \sum_{i=0}^{n-1} \hat{q}_i \frac{d^i x(t_j)}{dt^i},$$

where \hat{p}_i and \hat{q}_i represent the unknowns of this equation.

The required values of the derivatives can be determined by using analogue differentiators in conjunction with the analogue signals, or by sampling the signals and using numerical techniques. However, differentiation is a noisy process, and integration can be performed more easily and accurately than differentiation when working with either analogue or digital equipment. Thus it is desirable that an identification procedure employ integration rather than differentiation whenever possible.

Diamessis^{(45),(46)} considered the possibility of integrating an equation of the form of Equation 2.6 before establishing a system of simultaneous equations; however, his procedure depends upon knowing values of the derivatives of both the excitation and the response before the final system of equations can be established. Thus, his procedure still uses the less reliable method of differentiation. Diamessis did not consider the non-ideal case or the problem of implementing his procedure thus his procedure has not been completely investigated.

In the next section, a method will be presented, for establishing a system of equations, which employs only the process of integration. The established method will then be fully implemented and investigated in the remainder of this research.

Development of a System of Linear Equations

If Equation 2.6 is repeatedly integrated n times, the result is a time dependent equation which relates the response variable $x(t)$ to integrated values of $x(t)$, integrated values of $y(t)$, and a set of transformed initial conditions. A system of linear, algebraic equations can then be established based only upon measured values of the excitation and response of an unknown system. The solution of this system of equations will then yield the values of the coefficients and initial conditions of the differential equation which describes the unknown system. In the following discussion the symbols x and y will be used to denote the signals $x(t)$ and $y(t)$.

The desired integration is given by

$$\underbrace{\int_0^{t_m} \dots \int_0^{t_v}}_n \left[x^{(n)} + p_{n-1} x^{(n-1)} + \dots + p_0 x - q_{n-1} y^{(n-1)} - \dots - q_0 y \right] dt_v \dots dt_m = 0, \quad (2.9)$$

and the indicated integration must be performed with proper consideration given to the arbitrary initial conditions which can exist within the system. The integration of Equation 2.9 will now be carried out so as to illustrate the results of integrating the individual terms of this equation.

Integrating the n^{th} derivative of x on the interval $[0, t_v]$ yields

$$\underbrace{\int_0^{t_m} \dots \int_0^{t_u} \int_0^{t_v}}_n x^{(n)} dt_v dt_u \dots dt_m = \quad (2.10)$$

$$\underbrace{\int_0^{t_m} \dots \int_0^{t_u}}_{n-1} [x^{(n-1)} - x^{(n-1)}(0)] dt_u \dots dt_m.$$

Proceeding in this same manner, the final result of performing all integrations in Equation 2.10 is

$$\underbrace{\int_0^{t_m} \dots \int_0^{t_v}}_n x^{(n)} dt_v \dots dt_m = \quad (2.11)$$

$$x(t_m) - \sum_{i=0}^{n-1} x^{(i)}(0) \frac{(t_m)^i}{(i)!},$$

where $x^{(0)}(0) \equiv x(0)$.

In the same way, integration of the term involving the k^{th} derivative of x for $k = (0, 1, \dots, n-1)$ yields

$$p_k \underbrace{\int_0^{t_m} \dots \int_0^{t_v}}_n x^{(k)} dt_v \dots dt_m = \quad (2.12)$$

$$\underbrace{p_k \int_0^{t_m} \dots \int_0^{t_r}}_{n-k} x dt_r \dots dt_m = p_k \sum_{i=0}^{k-1} x^{(i)}(0) \frac{(t_m)^{n-k+i}}{(n-k+1)!},$$

where the summation is understood to exist only when $k-1 \geq i$. Similarly, integration of the terms involving the k^{th} derivative of $y(t)$ for $k = (0, 1, \dots, n-1)$ yields

$$q_k \underbrace{\int_0^{t_m} \dots \int_0^{t_v}}_n y^{(k)} dt_v \dots dt_m = \quad (2.13)$$

$$q_k \underbrace{\int_0^{t_m} \dots \int_0^{t_r}}_{n-k} y dt_r \dots dt_m = q_k \sum_{i=0}^{k-1} y^{(i)}(0) \frac{(t_m)^{n-k+i}}{(n-k+i)!}.$$

For notational purposes, the $(n-k)$ integrations of the variable $-x(t)$ on the interval $[0, t_m]$ will be given by

$$A_{n-k}(-x, t_m) = \underbrace{\int_0^{t_m} \dots \int_0^{t_r}}_{n-k} -x dt_r \dots dt_m, \quad (2.14)$$

for $k = (0, 1, \dots, n-1)$. Similarly, the $(n-k)$ integrations of the variable $y(t)$ will be given by

$$B_{n-k}(y, t_m) = \underbrace{\int_0^{t_m} \dots \int_0^{t_r}}_{n-k} y dt_r \dots dt_m. \quad (2.15)$$

Combining the results of Equations 2.11, 2.12, 2.13, and using the notational representations of Equations 2.14 and 2.15, it is possible to write the results of integrating Equation 2.9 as

$$x(t_m) = \sum_{i=0}^{n-1} x^{(i)}(0) \frac{(t_m)^i}{(i)!} - \sum_{i=0}^{n-1} p_i A_{n-i}(-x, t_m) \quad (2.16)$$

$$\begin{aligned} &= \sum_{i=0}^{n-1} q_i B_{n-i}(y, t_m) - \sum_{k=0}^{n-1} p_k \sum_{i=0}^{k-1} x^{(i)}(0) \frac{(t_m)^{n-k+i}}{(n-k+i)!} \\ &+ \sum_{k=0}^{n-1} q_k \sum_{i=0}^{k-1} y^{(i)}(0) \frac{(t_m)^{n-k+i}}{(n-k+i)!} = 0. \end{aligned}$$

This equation can be simplified by combining into a single term all expressions involving the i^{th} power of t_m . The simplified expression is given by

$$\begin{aligned} x(t_m) = & \sum_{i=0}^{n-1} p_i A_{n-i}(-x, t_m) + \sum_{i=0}^{n-1} q_i B_{n-i}(y, t_m) \\ & + \sum_{i=0}^{n-1} x_{i+1}^{(i)}(0) \frac{(t_m)^i}{(i)!}. \end{aligned} \quad (2.17)$$

In Appendix A, it is shown that the initial conditions of this equation, now denoted as $x_i(0)$, are the initial conditions of an equivalent first order vector differential equation representation of Equation 2.6. The transformation between the initial conditions of the vector differential equation and those of the scalar differential equation is derived in Appendix A, and is given as Equation A.35.

Equation 2.17 is valid at any time t_m , and the operators $A_k(-x, t_m)$ and $B_k(y, t_m)$ indicate the results which would be obtained if an analogue

signal is integrated k times by analogue integrators, or if a sampled signal is integrated k times by numerical methods. Thus, all of the quantities of Equation 2.17 which depend upon time can be determined from measured values of $x(t)$ and $y(t)$. It is then theoretically possible to evaluate the time dependent quantities of Equation 2.17, at enough different values of time, to establish a system of linear algebraic equations. The unknowns of this system of equations are then the parameters of a model.

Defining the Identification Procedure

The problem of identifying a system has now been reduced to the problem of establishing and solving a system of linear, algebraic equations. The identification process is defined as the process by which measured data is operated upon so as to produce a model. The data presented to the identification procedure can then take two forms:

1. The excitation (when the excitation is other than an unit impulse) and response of a system are sampled. The sampled data serves as input to the identification procedure, and is to be integrated numerically within the procedure.

2. The excitation (when the excitation is other than an unit impulse) and response of a system are integrated with analogue integrators. The required measurements are made, and serve as input to the identification procedure.

The required measurements will be specified during the development of the identification procedure. Chapter VII also contains specifications upon the excitation for the case where the excitation is other than an unit impulse.

Formation of the Normal Equations

If an unknown system is excited by a signal $y(t)$, and the excitation and response along with integrated values of these signals are measured at T discrete times, it is then possible to form T equations of the form of Equation 2.17. This system of equations will have $3 \times s$ unknowns, where the value of s is as yet unspecified; however, it will be required that enough measurements are taken so that $T \geq 3 \times s$, for a specified value of s . Such an established system of linear equations will have the form

$$A \underline{c} = \underline{b}, \quad (2.18)$$

where

A is $(T \times k)$ and known,

\underline{b} is $(T \times 1)$ and known,

\underline{c} is $(k \times 1)$ and unknown, and

$$T \geq k = 3 \times s.$$

If $\hat{\underline{c}}$ is a possible solution of this system of equations, then a residue vector \underline{r} is defined by

$$\underline{r} = \underline{b} - A \hat{\underline{c}}. \quad (2.19)$$

where \underline{r} is $(T \times 1)$. A norm of the residue vector is defined as the sum of the squares of the elements of \underline{r} , and the value of the norm is zero when Equation 2.18 represents a system of equations for which $\hat{\underline{c}}$ is an exact solution.

When a system of equations is given in the form of Equation 2.18, it is desirable to obtain a solution which will yield the minimum norm for the residue vector. This solution is referred to as the least squares solution, and can be obtained by establishing and solving the normal equations^{*}. The normal equations are formed by premultiplying both sides of Equation 2.18 by the transpose of A , and are then given by

$$(A'A) \underline{c} = (A'\underline{b}), \quad (2.20)$$

where

$A'A$ is $(k \times k)$ and known,

$A'\underline{b}$ is $(k \times 1)$ and known, and

\underline{c} is $(k \times 1)$ and unknown.

A standard method of solving a system of linear, algebraic equations can now be employed to find the least squares solution.

It will be noted that the formation of the normal equations allows a system of equations to be established in which the number of equations can exceed the number of unknowns. This is done with the assurance that the solution will be the least squares solution for all equations. For the ideal case, only as many equations will be required as there are unknowns since the least squares solution for a system of $(k \times k)$ equations will result in a zero residue vector. This solution will then be the exact solution for all other equations which could have been used.

^{*}Householder (26), p. 72; Whittaker (27).

The use of the normal equations allows the response of the model to be matched to the response of the system at as many sample points as desired. This is particularly important in the non-ideal case. For the non-ideal case, the residue vector will contain some non-zero elements, and the least squares solution must be found since all equations will not have the same solution. The error function ERMS is related to the norm of the residue vector, and minimizing the norm of \underline{r} also tends to minimize ERMS. This point will be discussed more in a later section.

The process of establishing and solving the required system of equations is by far the most difficult problem in actually implementing the identification procedure. In the remainder of this chapter, as well as in the next chapter, it will be assumed that the proper number of equations can be established and can be correctly solved. Chapters IV and V are then devoted to the actual problem of establishing and solving the required system of equations.

Determination of Order

If the unknown system is of order n , and a model is to be established of order s , then Equation 2.17 will have the form

$$\begin{aligned}
 x(t_m) = & \sum_{i=0}^{s-1} z_{i+1}(0) \frac{(t_m)^i}{(i)!} + \sum_{i=0}^{s-1} u_i A_{s-i}(-x, t_m) \\
 & + \sum_{i=0}^{s-1} v_i B_{s-i}(y, t_m).
 \end{aligned} \tag{2.21}$$

Here, $z_{i+1}(0)$, u_i , and v_i represents the unknown parameters of the model, and are to be determined by obtaining the least squares solution.

Three distinct cases can occur when the value of s is arbitrarily established namely: $s \equiv n$, $s > n$, and $s < n$. These three cases are now discussed.

Case No. 1: $s \equiv n$

If the system of equations is established for the correct number of unknowns, then denoting the least squares solution by $\hat{z}_{i+1}(0)$, \hat{u}_i , and \hat{v}_i for $i = (1, 2, \dots, n-1)$, the j^{th} element of the residue vector will be given by

$$r(t_j) = x(t_j) - \sum_{i=0}^{n-1} \hat{z}_{i+1}(0) \frac{(t_j)^i}{(i)!} \quad (2.22)$$

$$= \sum_{i=0}^{n-1} \hat{u}_i A_{n-i}(-x, t_j) - \sum_{i=0}^{n-1} \hat{v}_i B_{n-i}(y, t_j).$$

If $x(t_j)$ is now replaced by its representation as given by Equation 2.17, then Equation 2.22 can be rewritten as

$$r(t_j) = \sum_{i=0}^{n-1} (x_{i+1}(0) - \hat{z}_{i+1}(0)) \frac{(t_j)^i}{(i)!} \quad (2.23)$$

$$+ \sum_{i=0}^{n-1} (p_i - \hat{u}_i) A_{n-i}(-x, t_j) + \sum_{i=0}^{n-1} (q_i - \hat{v}_i) B_{n-i}(y, t_j).$$

Since the functions of time of this equation are general weighting functions, it is valid to rewrite the equation as

$$r(t_j) = \sum_{i=1}^{3n} (c_i - \hat{c}_i) w_i(t_j). \quad (2.24)$$

Here, c_i is an element of \underline{c} , a vector representation of the parameters of the system. The values of \underline{c} are fixed but unknown. Similarly, $\hat{\underline{c}}$ represents the corresponding parameters of the model which is to be identified with the unknown system. The general weighting function $w_i(t_j)$ is fixed for a given problem.

It is now necessary to determine the parameters $\hat{\underline{c}}$ which will minimize the norm of the residue vector. From Equation 2.24, the square of the j^{th} element is given by

$$\begin{aligned} r(t_j)^2 = & \sum_{i=1}^{3n} (c_i - \hat{c}_i)^2 w_i^2(t_j) \\ & + 2 \sum_{i=1}^{3n} \sum_{k=i+1}^{3n} (c_i - \hat{c}_i)(c_k - \hat{c}_k) w_i(t_j) w_k(t_j), \end{aligned} \quad (2.25)$$

and its partial derivative with respect to \hat{c}_e , for $e = (1, \dots, 3n)$, is

$$\frac{\partial r(t_j)^2}{\partial \hat{c}_e} = -2 \sum_{i=1}^{3n} (c_i - \hat{c}_i) w_e(t_j) w_i(t_j). \quad (2.26)$$

The only general solution for which all partial derivatives of all

squared elements of the residue vector are zero thus occurs when

$$\hat{\underline{c}} \equiv \underline{c} , \quad (2.27)$$

and for this solution, the norm of the residue vector is identically zero. This then corresponds to exact identification of the system, hence, $E(t;y)$ is identically zero for all time.

It is now been determined that when the system of linear algebraic equations are established for a model of the same order as the system being identified, then the least squares solution will yield a model which is identically equal to the system. It should be noted that this discussion has been for the ideal case where the data is exact, and exact numerical solutions can be obtained. The non-ideal case will be treated in the next chapter, and it will then be shown that the errors of the solution will be bounded by an expression involving the errors in the established system of equations.

Case No. 2: $s > n$

If the chosen value of s is larger than the actual order of the system, the solution of the established system of algebraic equations, if it exists, will yield a model of order s . The parameters of the model will yield the minimum residue vector. This minimum value will be zero only if the solution is exact; however, a norm of zero value occurs if and only if the response of the model and the response of the system are identical, for the given excitation. It will now be shown that for $s > n$ the established system of equations has infinitely many solutions which result in a minimum residue vector; therefore, the coefficient matrix is singular.

Assume that a system of order n is to be identified, and let its transfer function be denoted by $H_n(S)$. The system is now excited, and the proper measurements are made in order to establish the required system of equations. The equations are established for a model of order s , where the value of s is arbitrary, but greater than n . Consider now a second system of order s , and let its transfer function be denoted by $H_s(S)$. If $H_s(S)$ is formed by multiplying $H_n(S)$ by a transfer function $H(S)$, where $H(S)$ contains only surplus factors*, then the response of $H_s(S)$ will be identical to the response of $H_n(S)$ for all excitations. The parameters of $H_s(S)$ will then be an exact solution of the established system of equations.

The choice of the surplus factors of $H(S)$ is arbitrary, and it is possible to form infinitely many variations, therefore, infinitely many systems of the type $H_s(S)$ exist whose parameters are exact solutions to the system of equations. Since it is impossible to find an unique solution, this requires that the coefficient matrix of the system of equations be singular. Based upon this, it is now possible to state that the maximum rank of a system of equations, established in connection with the identification procedure of this research, corresponds to the order of the system of equations established for case no. one. The above situation of having a singular coefficient matrix corresponds to the situation that exists whenever a system of equations is established in which some of the equations are dependent.

If the order of the system is unknown, it is possible to establish a system of equations for several different order models. The

* Surplus factors are roots that are common to both the numerator and denominator polynomials of the transfer function.

largest non-singular system of equations will then yield the order of the system, and the solution of this system of equations will yield the parameters of the model. Again, it should be noted that this discussion is for the ideal situation; however, the non-ideal situation will be treated in the later chapters.

Case No. 3: $s < n$

An attempt at solving the identification procedure for $s < n$ is an attempt at finding a model of lower order than the true system such that the model exhibits as close as possible the same response characteristics as the true system. The closeness by which a given system can be approximated by a lower order model will depend upon the characteristics of the system as well as the difference in orders between the system and the model. It will be stated now that the identification procedure can be solved not only for a true model, but also for lower order models which approximates the system. This aspect of the identification procedure will be discussed more in the following chapters.

Response Error and the Least Squares Solution

Although the identification procedure does not directly attempt to determine a model which minimizes the response error ERMS, it does attempt this minimization in an indirect way. This can be seen by examining the relationship between the instantaneous error $E(t_j; y)$ and the element of the residue vector for an equation formed at the time t_j .

Assume that the response and excitation of a system has been measured and that a system of $(T \times k)$ equations of the form

$$A_S \underline{c} = \underline{b} , \quad (2.28)$$

has been established for $T \geq k$. Here, A_S denotes that the coefficient matrix has been established from measurements made on the system. If $\hat{\underline{c}}$ is the least squares solution for these equations then the residue vector is given by

$$\underline{r}_S = \underline{b} - A_S \hat{\underline{c}} \quad (2.29)$$

where the j^{th} element of this system of equations has the form of Equation 2.22.

Since $\hat{\underline{c}}$ represents the k parameters of a model of order s , it is now possible to compute the response of this model for the same excitation $y(t)$, and develop a system of equations of the form

$$A_M \underline{c} = \underline{b} , \quad (2.30)$$

where \underline{b} is the vector of Equation 2.28, and A_M is formed from measurements made on the excitation and response of the model. The individual equations of this system of equations are established at the same discrete times t_j as the corresponding equations of Equation 2.28. A residue vector is now defined as

$$\underline{r}_M = \underline{b} - A_M \hat{\underline{c}} \quad (2.31)$$

where $\hat{\underline{c}}$ is the least squares solution of the original system of equations.

If the instantaneous error $E(t_j; y)$ is computed using the measured responses of the system and the model at every instantaneous time t_j

for which an equation in the above system of equations has been established, it is seen that

$$r_{M_j} = E(t_j; y), \quad (2.32)$$

for every value of j . However, the j^{th} element of the residue vector \underline{r}_S is identically equal to r_{M_j} only if

$$A_S \equiv A_M. \quad (2.33)$$

This corresponds to the exact case, and for this case the response error of the model and the system is identically zero, and

$$r_{S_j} = r_{M_j} = E(t_j; y) = 0. \quad (2.34)$$

For the inexact case the norm of \underline{r}_S will not necessarily be zero although $\hat{\underline{c}}$ is the least squares solution. This then implies that the element of $\hat{\underline{c}}$ are not the exact parameters of the system; thus, the error function ERMS will not be identically zero. It is seen that if A_S is close to being exact, then the least squares solution will be close to the value of the true parameters of the system. It is then necessary to determine the uncertainty in the solution of an established system of equations, based upon a knowledge of the uncertainties in the established system of equations. A part of the next chapter is devoted to this problem.

Class of Allowable Excitation

This research is concerned mainly with identification when the excitation is an unit impulse; thus, Equation 2.17 must be considered

for the situation when $y(t)$ is an impulse. Consideration must also be given to the uniqueness of a model in relationship to the excitation used for the purpose of identification whenever the excitation is other than an unit impulse.

Excitation is an Unit Impulse

If $y(t)$ is an unit impulse, its effect is to change the existing initial conditions within the system. Mathematically the operator $B_{n-i}(y, t_m)$ of Equation 2.17 becomes an expression involving only time, and is given by

$$B_{n-i}(\delta(t), t_m) = \frac{(t_m)^{n-i}}{(n-i)!}, \quad (2.35)$$

where $\delta(t)$ represents an unit impulse. With this condition, Equation 2.17 now has the form

$$x(t_m) = \sum_{i=0}^{n-1} \frac{c_i (t_m)^i}{(i)!} + \sum_{i=0}^{n-1} p_i A_{n-i}(-x, t_m), \quad (2.36)$$

where $c_i = x_{i+1}(0) + q_{n-i-1}$.

The system of equations, established for the identification procedure, now has only $2 \times n$ unknowns since it is impossible to distinguish between the existing initial conditions and the new initial after the application of the impulse. If the system has zero initial conditions prior to its excitation by an unit impulse, the $2 \times n$ unknowns are identically the values of the coefficients of the differential equation which describes the unknown system. If identification

is attempted when the response of the system is due to arbitrary initial conditions, the results will be a model which has the measured response as its impulse response. In this way, it is possible to find an expression of time which fits the given response. This method of identification has application as will be illustrated in Chapter VI.

If a system is represented by a model such that the model and the system have the same impulse response, then the system has been uniquely identified. That is, the responses of the system and the model will be identical for any excitation. However, if the impulse response of the model only approximates the impulse response of the system (i.e. $ERMS \neq 0$), then the user must determine the possible effect this can have in relation to the intended use of the model. Sarber⁽²⁴⁾ gives a different performance index from that given by Equation 2.6, and the reader is referenced to that paper for another method for evaluating the accuracy and closeness of a model derived based upon the measured impulse responses of the system.

Excitation is Other Than an Unit Impulse

When a system is excited by an input other than an unit impulse, it is possible to separate all coefficients and initial conditions of Equation 2.17. If the initial conditions are known prior to the start of the identification process, it is possible to reduce the system of equations to only $2 \times n$ unknowns; otherwise, the value of the initial conditions must be carried as unknowns.

There are requirements upon the excitation used to excite the system. The first requirement, from a mathematical standpoint, is that the excitation must not contain impulses. In this way, the values of

the initial conditions cannot change instantaneously, and can be carried as constants in the system of linear equations. A second requirement upon the excitation has to do with uniqueness of identification.

If identification is to be attempted with an arbitrary excitation, it is necessary that the response of the system be complete for this excitation. That is for example, identification cannot be guaranteed successful when the excitation is a pure sine function of fixed frequency. While identification can be attempted with such an excitation, there are no guarantees that a model established with an arbitrary excitation will have a response approximating that of the system when the excitation is changed. Since the excitation must completely excite the system, it will be impossible in general to correctly identify a system with data gathered during normal operating time. This problem will be discussed in Chapter VII.

CHAPTER III

ERROR ANALYSIS

This chapter contains an error analysis of the developed identification procedure. All computational sources of error for an identification problem are indicated, and consideration is given to the method for estimating the uncertainties of the parameters determined for the model. The total uncertainty is due to inexactness in both the required data and the numerical methods employed to solve the identification problem. This chapter is applicable for the general identification problem -- that is when the system has been excited with a general signal $y(t)$.

Sources of Error

Figure 2 shows in block diagram form the sources of error for the identification procedure developed in the preceding chapter. Sources of error which effects the data are indicated with dashed lines, while sources of error due to the numerical methods, employed by the actual identification procedure, are indicated in solid lines. There are two numerical methods required within the identification procedure, and these are methods for

1. numerically integrating a sampled signal,
2. numerically solving a system of linear, algebraic equations.

The object of this chapter is, then, to indicate the sources of error, and to discuss how the errors which occur within the actual

identification process can be minimized. Methods will also be given by which the uncertainty of the solution of an identification problem can be estimated. This will require knowledge as to the uncertainty of the available data as well as knowledge as to the nature of the errors which occur within the identification procedure. Each type of error will be given a symbol, and these symbols are

ε_d - error in data,

ε_I - error due to numerical integration methods,

ε_s - error in the solution of a system of linear,
algebraic equations.

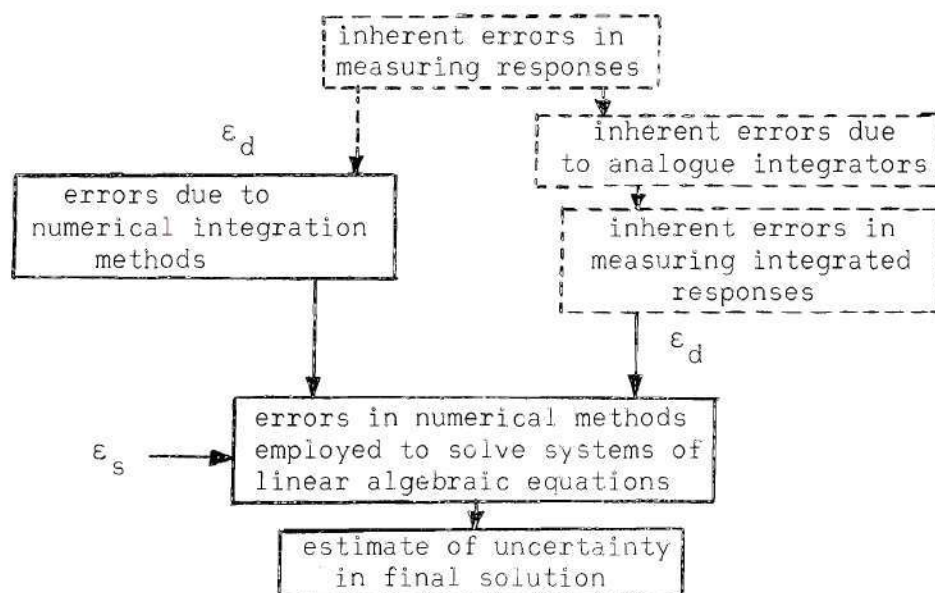


Figure 2. Sources of Errors.

Inherent Errors

Inherent errors are errors in the value of the data caused by uncertainty in measurements, or by the necessarily approximate nature of representing the measurement in a finite number of digits. The effect of additive noise is also an inherent error. Except for the case of unrealistically small assumed noise levels, it is essential to use a reasonable degree of redundancy* in the observed operating records in order to keep the noise-influenced fluctuation in parameter estimates within tolerable bounds.

Even if a true signal is integrated by analogue integrators, there still exist inherent errors in the integrated signal because of the non-ideal characteristics of analogue integrators. This is especially true at low frequencies where the integrators introduces a phase shift as well as a non-ideal magnitude characteristic**.

In general, there will be available some estimate of the accuracy of the measured data. This will usually take the form of specifying that the data is correct to "d" significant figures, or by specifying the possible inaccuracies in the data. Based upon this knowledge, it will be possible to make a final estimate as to the uncertainty in the computed values of the parameters of the model.

Errors in Numerical Integration Methods

There are two sources of error when a signal is integrated by numerical methods. These errors are first truncation error due to the

* By redundancy, in this case, is meant the averaging of a finite number of samples in order to obtain an average value.

** Johnson (39), pp. 189-191.

approximate nature of the integration methods, and second roundoff error due to the arithmetic operations involved. Estimates of the bounds for each of these errors can be calculated, and it will be shown that the sampling interval must be chosen so as to minimize the total effect of these errors.

Truncation Error

If a signal is integrated by numerical methods, there exists a truncation error due to the fact that the true curve is fitted between sample values by an approximate curve when computing the area of integration. Any of the several Newton-Cotes integration formulas* can be used in conjunction with the identification procedure, and it is possible to estimate a bound on the size of truncation error for each method.

It is desirable to use a numerical integration procedure which preserves the number of available samples since the i^{th} integration must be calculated numerically from the $(i-1)^{\text{th}}$ integrated values. The trapezoidal rule possesses this sample-preserving characteristic; however, the higher order methods, for integrating numerically, can be used in conjunction with the trapezoidal rule to obtain the desired results. For example, if Simpson's rule is used it is possible to obtain a value of the integrated function at every odd sample of the original data. If the trapezoidal rule is used to obtain a value of the integrated function at the second sample, Simpson's rule can then be used to obtain an integrated value at each of the remaining even sample points. Similarly, if an n^{th} order Newton-Cotes formula is used, the

*Kunz(40), pp. 145-147.

trapezoidal rule must be used to obtain integrated values for $(n-2)$ points.

It is desirable to use the higher order numerical methods because the truncation error will, in general, become smaller as the order of the numerical method increases. Probably the most widely known technique for integrating a sampled function is Simpson's rule. It is surprisingly accurate, and there is some mathematical basis for a general statement that Simpson's rule with k points provides the same general order of accuracy as the trapezoidal rule does with $2 \times k$ points*.

The truncation errors resulting when a function is numerically integrated by either the trapezoidal rule or Simpson's rule are given respectively by

$$E_{TR} = \frac{1}{12} T H^2 \frac{d^2 f(c)}{dt^2}, \quad (3.1)$$

and

$$E_{SR} = \frac{1}{180} T H^4 \frac{d^4 f(c)}{dt^4} \quad (3.2)$$

where

T - time interval of integration,

H - time interval between samples,

$\frac{d^n f(c)}{dt^n}$ - n^{th} derivative of the function evaluated at point c within the interval of integration.

The point c where the derivatives are evaluated is given by the mean

* McCracken (43), p. 179.

value theorem; however, as an upper bound on the size of truncation error, the respective derivatives can be evaluated at their maximum values.

The error equations for the other Newton-Cotes formulars are given by Kunz ⁽⁴⁰⁾, and have the same form as Equations 3.2 and 3.3. In all cases, once the function to be integrated and the desired time interval T has been specified, the only variable is the sampling interval H. Thus decreasing the value of H will decrease the size of the truncation error in all cases.

Roundoff Error

Roundoff errors occur during any arithmetic operation on a digital computer, because the computer must work with a fixed number of digits. If a sampled signal has an absolute error, at each sampled value, less than 10^{-d} , where d represents the number of significant figures, it is possible to estimate a bound on the rounding error due to the arithmetic processes of a numerical integration method. This bound* is given by

$$|e_r| \leq |\bar{y}| \frac{5 \times 10^{-d}}{2} T^2 H^{-p}, \quad (3.3)$$

where

e_r - error due to roundoff,

\bar{y} - average value of function being integrated,

T - time length of required integration,

H - interval between samples,

* McCracken (43), pp. 166-171.

d = number of significant figures of data,

$$0 \leq p \leq 1.$$

The value of p is not known exactly, but its value lies within the indicated range. This equation is approximately true for numerical integration by both Simpson's rule and the trapezoidal rule.

It must be noted that there are several assumptions made during the derivation of Equation 3.3 that must be satisfied before the calculated bound can be completely valid. These assumptions are given by McCracken* and will not be repeated here since the equation is repeated only to indicate what factors effect round-off error.

In Equation 3.3, d actually represents the number of digits used in the arithmetic process; therefore, it is assumed that the data is given to d significant figures, and all arithmetic operations are also rounded to d significant figures. In actual practice, it will generally be true that the data will be accurate to d significant figures while the arithmetic operations will be performed with $k > d$ digits. For example, the computer used in this research is the B5500 which has a word length of approximately eleven digits. In general the data will never be more accurate than four or five digits; thus, rounding errors will not be as serious as Equation 3.3 indicates.

Comparison of Equations 3.1, 3.2, and 3.3 indicates a discrepancy between theoretical and practical computations. In theory, the truncation error can be made as small as desired by taking H significantly small. In practice, however, roundoff error prevents all numerical methods from obtaining this arbitrarily small inaccuracy.

* McCracken (43), pp. 166-171.

Thus, when using any numerical method for integrating a given function, it is necessary to choose the sampling interval so as to minimize the total error due to both truncation and roundoff errors.

Since the total error, due to numerical integration, depends upon the function being integrated as well as the sampling interval, it is impossible to establish explicit rules as to the size of the sampling interval to be used for each situation. Hopefully, an estimate based on prior experience will serve as a guide; however, in any case it is possible to estimate bounds on the individual errors once a sampled signal is available. It can then be determined if the sampling interval must be changed and/or the number of significant figures in the available data increased.

Total Bound on Errors

The final bound on the errors due to numerical integration is given by

$$\varepsilon_I \leq |e_r| + |E_{NI}|, \quad (3.4)$$

where E_{NI} represents the truncation error due to the particular numerical rule that is used. The total inaccuracy of the data, after the first integration is then bounded by

$$\varepsilon_I^1 \leq \varepsilon_I + |\varepsilon_d| \quad (3.5)$$

where the superscript 1 denotes that this is the first integration.

In each of the following integrations, it is necessary to first compute the bound ε_I for the function being integrated, and then the total

bound is given by

$$\epsilon_I^J \leq \epsilon_I^{J-1} + \epsilon_I^J, \quad (3.6)$$

for $J = (2, 3, \dots, s)$. It must be realized that when errors are added, proper consideration must be given to the data for which the errors are computed. That is, if the sampled signal has only d significant figures, then Equation 3.5 indicates that the integrated signal can have no more than d significant figures, and in general will have less than d significant figures because of truncation and roundoff errors.

Roundoff Errors and Ill-Conditioned Matrices

A final source of error in the solution of the identification problem occurs during the actual solving of the established system of algebraic equations. Even if all the data is exact, roundoff errors in simple arithmetic processes as well as the existence of ill-conditioned coefficient matrices can lead to solutions with little reliability.

The effect of an ill-conditioned matrix can be illustrated as follows. Assume a system of linear algebraic equations of the form

$$A \underline{c} = \underline{b}, \quad (3.7)$$

is to be solved. Let the true solution be denoted by \underline{c}' , and let an available solution be denoted by $\hat{\underline{c}}$. The error vector of the solution is then given by

$$\underline{e} = \underline{c}' - \hat{\underline{c}} = A^{-1} \underline{b} - \hat{\underline{c}}, \quad (3.8)$$

where A^{-1} is the true inverse of the A matrix. Similarly, the residue vector is given by

$$\underline{r} = \underline{b} - A \hat{\underline{c}}. \quad (3.9)$$

It is hoped that $\underline{e} = \underline{r} = 0$; however, because of inexact data, it is generally impossible to obtain such accuracy in the solution. However, it is possible to relate the error and residue vectors by combining Equations 3.8 and 3.9. This relation is given by

$$\underline{e} = A^{-1} \underline{r}. \quad (3.10)$$

If $\hat{\underline{c}}$ is not the true solution, the residue vector will not be zero; however, even if the residue vector contains small element values, its errors are reflected into the error vector by the size of the elements of the inverse A matrix. If the elements of A^{-1} are large, the error vector can then have large element values although the residue vector has relatively small elements. In general, an established system of equations will not have an exact solution; therefore, it is necessary to find a least squares solution. The uncertainty of accepting a least squares solution lies in the fact that a system of equations can have several solutions, all of which have a relatively small residue vector; however, Equation 3.10 shows that all of these solutions can be far removed from the true solution. This general problem is referred to as the problem of ill-conditioned matrices, and the resulting practical problem will be discussed in more detail in the following chapters.

Ill-conditioning of a system of equations tends to increase the

effect of roundoff errors which occur during the process of solving a system of simultaneous equations. In solving the established system of equations, it is then necessary to seek rules which will hold the round-off errors as small as possible. There are methods of achieving this result, and these methods can be applied separately.

A method employed during the process of solving a system of equations, to reduce rounding error, is called "pivotal consideration." Pivoting refers to the process of interchanging rows during the application of a direct method, such as Gauss elimination, for solving a system of equations*. The interchanging of rows is necessary in order to work with the largest number (in absolute value), during the elimination process. In a strict sense, it is necessary to employ pivoting both with respect to rows and columns; however, this increases the computation time involved.

After a solution has been obtained, another technique can be employed to refine this solution. The refining technique has the form of an iterative procedure, and can permit a reasonable solution of some ill-conditioned systems of equations. One iterative scheme for refining a solution obtained by Gauss elimination is presented by McCracken⁽⁴⁴⁾, and a second method for improving an approximate inverse of a matrix is given by Rodman⁽³³⁾.

The effect of roundoff error, while solving a system of equations, can generally be detected and removed from the final solution. It is thus necessary that any implementation of the identification procedure include methods for refining a solution.

* McCracken (44), pp. 238-246.

Uncertainty in the Parameters

Sources of computational errors have been indicated, and discussions have been given as to the minimization of the errors which occur within the identification procedure. It is now necessary to consider the final solution, and estimate the uncertainty of each parameter of the model. If the data is exact, the established system of linear, algebraic equations can be represented in notational form by

$$A \underline{c} = \underline{b}, \quad (3.11)$$

and the exact solution is given by \underline{c}' . However, due to inherent errors as well as the errors due to numerical integration, the actual established system of equations is given by

$$(A+E)(\underline{c}+h) = (\underline{b}+\underline{k}), \quad (3.12)$$

where it is impossible to physically separate the quantities within the parenthesis. The least squares solution of this equation can be denoted by $\hat{\underline{c}}$, where $\hat{\underline{c}} = \underline{c}' + \hat{\underline{h}}$.

It will be noted that the exact solution \underline{c}' represents the true parameters of the unknown system, while the least squares solution $\hat{\underline{c}}$ represents the parameters of the resulting model. Thus, $\hat{\underline{h}}$ represents the error between the parameters of the true system, and the parameters of the model. It is thus desired to determine a bound on the size of the elements of $\hat{\underline{h}}$, knowing the bounds on the elements of the error matrix E , and the error vector \underline{k} .

It is possible to compute bounds for the size of the elements of E and \underline{k} based upon the previous discussions. All of the elements of \underline{k}

are simply the sampled values of the response of the system, thus each element is bounded in value by ϵ_d . Similarly, an element of E is bounded by ϵ_I^j , where ϵ_I^j is calculated depending upon whether the corresponding element of the A matrix represents either a value of the operator $A_j(-x, t)$ or $B_j(y, t)$. The final results will be a bound on the errors, and will give an indication as to the number of significant figures contained in the coefficients and right hand members of the established system of equations.

Appendix C contains an analysis of Equation 3.12 in terms of the norms of vectors and matrices, and a bound is established for the relative error of the solution in terms of the relative errors in the established system of equations. The final result is given as Equation C.33, and is

$$\frac{\frac{||\hat{h}||}{||\underline{c}'||}}{1 - ||A|| ||A^{-1}|| \frac{||E||}{||A||}} = \frac{||A|| ||A^{-1}|| \frac{||\underline{k}||}{||\underline{b}||}}{1 - ||A|| ||A^{-1}|| \frac{||E||}{||A||}} + \frac{||A|| ||A^{-1}|| \frac{||E||}{||A||}}{1 - ||A|| ||A^{-1}|| \frac{||E||}{||A||}} \quad (3.13)$$

It is impossible to compute this bound since its computation depends upon knowing the actual errors; however, this expression clearly indicates the factors that effect the accuracy of a possible solution.

The first requirement for a reliable solution is that the relative errors of the system of equations be small, and the second requirement is that the quantity $\|A\| \|A^{-1}\|$ be small. An indication of the "condition" of a system of equations is defined by Equation C.34 as

$$\eta = \|A\| \|A^{-1}\| ,$$

where η is called the spectral conditioning number. If the value of η is relatively large then the system of equations is said to be ill-condition. The vector and matrix norms used in Equation 3.13 are discussed in Appendix C.

A method is now presented by which a measure of the uncertainty in the solution can be calculated. The derivation is given in Hildebrand⁽⁴⁷⁾, and only the results will be given here. It is now assumed that the coefficients and right hand members of the system of equations are known to d significant figures. A solution of Equation 3.12 is obtained by matrix inversion, and this solution is denoted by \hat{c} . If \tilde{a}_{ij} denotes the elements of the computed inverse of the $(A + E)$ matrix, then the uncertainty of the parameter \hat{c}_i is denoted by $\pm \delta \hat{c}_i$, where

$$\delta \hat{c}_i < (|\tilde{a}_{i1}| + |\tilde{a}_{i2}| + \dots + |\tilde{a}_{iN}|)D. \quad (3.14)$$

Here, the system of equations is $(N \times N)$, and the summation is taken over the i^{th} row of the inverse matrix. The quantity denoted by D is

$$D = (1 + |\hat{c}_1| + |\hat{c}_2| + \dots + |\hat{c}_N|) \frac{10^{-9}}{2} \quad (3.15)$$

A second method for obtaining estimates of the uncertainty in the parameters is given by McCracken⁽⁴⁴⁾. This method does not require

knowledge of the inverse of the coefficient matrix, and can be used in conjunction with an elimination method for obtaining the solution of the system of equations. The method will not be given here; however, the method does give an estimate of the uncertainty of each coefficient. It will be mentioned that McCracken's method gives extremely conservative bounds in that they are always greater than or equal to the bounds obtained by the method of Hilderbrand as given above.

Examination of Equation 3.14 shows again the effect of ill-conditioning. If the elements of the inverse matrix are large, the resulting uncertainties in the solution will be large. In order to reduce this range of uncertainty, it is thus necessary to carry a large number of significant figures in the given data. When this is not possible, it is then necessary to accept the fact that existing errors in the given data then permits the solution to be determined only within relatively wide error limits.

CHAPTER IV

GENERAL CONSIDERATIONS IN THE IMPLEMENTATION OF THE IDENTIFICATION PROCEDURE

Specification of the computerized realization of the theoretical identification procedure, developed in Chapter II, as well as specification of how the identification procedure is to be used in solving a practical identification problem is herein referred to as "implementation of the identification procedure." This chapter is concerned with three important aspects of the implementation problem, namely, the solving of a system of equations of the form of Equation 2.17 for the least squares solution, the determination of the uncertainty of an obtained solution, and the determination of the correct order of a model. While the discussion given in this chapter is for the situation where the system excitation is an unit impulse, the results are equally applicable for the situation where the system excitation is a general input.

Introduction

The implementation of the identification procedure has been conducted for three reasons:

1. To automate the basic identification procedure.
2. To investigate several aspects of the identification procedure.
3. To carry out the details of the identification procedure when given the appropriate data describing the response and excitation of an unknown system.

Before the identification procedure can be automated it is necessary to specify:

1. The manner in which the required system of simultaneous equations are to be established by specifying the individual times at which the equations are to be selected.
2. The manner that the established system of equations are to be solved for the least squares solution.
3. A method by which the identification procedure is to determine the correct order of the model.
4. The manner in which the uncertainty of the obtained solution is to be estimated.

Since it is necessary to work in practice with noisy data the choice of methods used in many of the steps of the automation must be based on empirical rather than completely theoretical studies.

The procedure used for investigating the automation aspect of the identification procedure is as follows:

1. Test data is obtained by sampling the impulse response of a known system.
2. The samples are integrated by a numerical method, and the accuracy of the integrated data is determined. This complete set of data serves as input to the identification procedure.
3. Several different approaches to establishing and solving the required system of equations are investigated.
4. The solutions obtained are compared to the known true solution in order to judge the reliability of each possible method for obtaining a solution.

The above procedure is repeated using a variety of known systems. Also, since the accuracy of the data in step one can be controlled, it is possible to simulate work with inaccurate data by contaminating the response of the test system with various levels of additive noise.

In giving the final way in which the identification procedure is to be implemented as well as the results of an investigation of the procedure, it is felt to be better to separate the presentation into several divisions, although each division is not necessarily disjoint. In particular, the divisions within this chapter are:

1. The method in which test data is generated.
2. The errors introduced when the samples are integrated by a numerical method, and the reliability of the computed bounds on the errors.
3. The accepted method for solving a system of $n \times k$ (for $n \geq k$) equations for the least squares solution.
4. The reliability of computed bounds on the uncertainty of the final solution.
5. The method for determining the order of the unknown system.

Chapter V completes the total presentation for the situation where the system excitation is an unit impulse, while Chapter VII completes the presentation for the situation where the excitation is a general signal.

Obtaining Test Data

To have data available for testing different stages of the implementation, the impulse response of a known system is first sampled and then integrated by a numerical method in order to form a table of data. The samples are taken over a period of time greater than the total

essential time length^{*} of the impulse response of the system, and a complete table of data includes the individual times at which the samples are taken, the value of the samples, and integrated values of the response at each sample time. Using the table of data it is possible to form an equation of the form of Equation 2.36 at each sample time although all equations are not necessarily required when solving a practical identification problem.

The response of the test systems, due to their excitation by an unit impulse, is generated by using a sixth order Runge-Kutta numerical integration procedure⁽³¹⁾. The response is generated and numerically integrated on a B5500 computer^{**} which has a word length of approximately eleven digits. In order to more closely simulate empirical data, the sampled impulse response of the system is truncated to five octal digits before integrating numerically. Also, the results of each integration step is truncated to five octal digits before repeating the integration process. Five octal digits corresponds to approximately five or six decimal digits depending upon the size of the number being truncated.

A combined implementation and investigation has been conducted using several systems of different orders, and using various sampling intervals and integration methods. For continuity, one test example has been chosen to illustrate certain significant results, and results of

^{*}The impulse response of a system extends over the interval $(0, \infty)$. For the purpose of this research, the essential time length is defined as the interval $(0, T)$ where T is the value of time such that the absolute value of the impulse response of the system on the interval (T, ∞) is less than one per cent of the maximum absolute value of the response.

^{**}Computer facilities were available through the Rich Electronic Computer Center at the Georgia Institute of Technology.

attempts at identifying this test example will be used throughout the discussions of both this and the following chapter. The test system is a fourth order system, and it has the properties given in Table 1. Figure 3 is a plot of the impulse response of the test system.

Table 1. Characteristics of a Test System

differential equation:	
(4) x	(3) $+ 19$
(2) x	(1) $+ 118$
(3) x	(2) $+ 320$
(1) x	(3) $+ 400x$
(3) $=$	(2) y
(1) $+ 7$	(1) y
(2) $+ 16$	(1) y
(1) $+ 10y$	
impulse response	
$x(t) = 1.32e^{-10t} - 0.31e^{-5t} + 0.17 \cos(2t - 95.5)e^{-2t}$	
poles of system	zeros of system
$-2 \pm j2$	$-3 \pm j1$
-5	-1
-10	

Errors Introduced by Numerical Integration

It is now known that the sampled response of the test system has approximately five significant figures, and it is necessary to determine the number of significant figures in the integrated data. The total knowledge of the number of significant figures in the generated table of data will then allow the calculation of the uncertainty of the solution of the identification problem. A procedure has been given in Chapter III for determining the bounds on the errors introduced when a signal is integrated by a numerical method; however, it must be recognized that computed bounds on the errors introduced by numerical integration methods

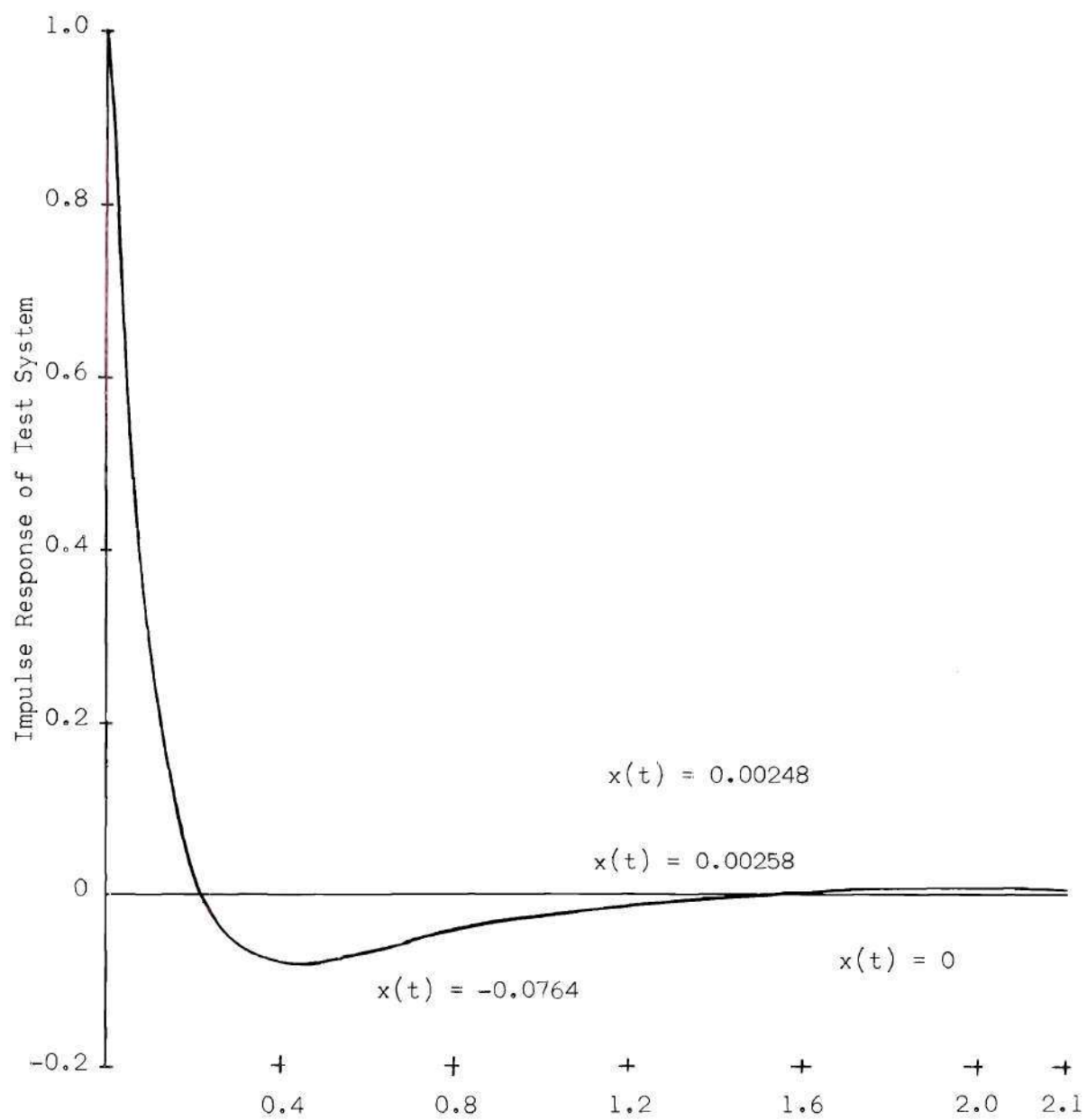


Figure 3. Plot of Impulse Response of Test System.

can in general give pessimistic results. The following discussion will illustrate this point as well as present a second method for estimating the uncertainty of the results of using a numerical integration process.

Estimating the Error Bounds Using Numerical Methods

The impulse response of the system of Table 1 has been generated for a total of two and one-quarter seconds, and it has been sampled with a sampling interval of 0.005 seconds. The signal is to be integrated using the trapezoidal rule, and from Inequality 3.3 it is possible to estimate the bound on the rounding error due to the integration process, while Equation 3.1 can be used to estimate the bound on truncation error. When computing these bounds, it is necessary to know both the average value of the function being integrated as well as the maximum value of its second derivative. The average value of the impulse response for the system of Table 1 is

$$\bar{x} = 0.01, \quad (4.1)$$

thus the bound on the roundoff error in the first integration process can be calculated as

$$|e_R| \leq 2.5 \times 10^{-4}. \quad (4.2)$$

The second derivative of the impulse function has the following characteristics

$$\left. \frac{d^2x}{dt^3} \right|_{\max.} = 125, \quad (4.3)$$

$$\left. \frac{d^2x}{dt^2} \right|_{\min.} = -0.16$$

$$\left. \frac{d^2x}{dt^2} \right|_{\text{ave.}} = 5.$$

The resulting bound on truncation error is

$$E_{\text{TR}} = 4.7 \times 10^{-6} \frac{d^2x(c)}{dt^2}, \quad (4.4)$$

where the value of the second derivative is in the interval $[-0.16, 125]$.

Using the largest possible value, the bound on the truncation error is

$$E_{\text{TR}} \leq 5.9 \times 10^{-4}. \quad (4.5)$$

Thus the total bound on the errors introduced by the numerical integration process, as given by Inequality 3.4 is

$$\epsilon_I \leq 8.4 \times 10^{-4}. \quad (4.6)$$

Using the trapezoidal rule to numerically integrate the sample impulse response over the interval $[0, 2.25]$ seconds, yields a value of 0.024170. Since the original data has only five significant figures, the integrated data can then have at most five significant figures.

Thus the bound ϵ_d is

$$\epsilon_d \leq 10^{-6}. \quad (4.7)$$

It is seen that ϵ_I dominates ϵ_d , and using Inequality 3.6 the final bound on integration error for the first order of integration is

$$\epsilon_I^1 \leq 8.4 \times 10^{-4} \quad (4.8)$$

This bound indicates that there is possible uncertainty in the third significant figure of the integrated data.

In computing the bounds on the error due to numerically integrating $x(t)$, it is seen that the large bound is due mainly to truncation and round-off error rather than ε_d ; however, it must be remembered that $|e_R|$ depends directly upon the number of significant figures of the function being integrated and that E_{TR} has been computed using a value of the second derivative that is approximately 25 times the size of the average value of the second derivative. It is necessary to use this large value of the second derivative; however, it can give a pessimistic result.

The integrated values of $x(t)$ form the sampled values of $A_1(x,t)$, and it is necessary to repeat the above process in order to determine a bound on the errors introduced when $A_1(x,t)$ is integrated to form $A_2(x,t)$. Upon inspecting the function which is to be integrated it is found that

$$|\overline{A_1(x,t)}| \approx 0.03, \quad (4.9)$$

$$\left| \frac{d^2 A_1(x,t)}{dt^2} \right|_{\max} \approx +12,$$

The value of the second derivative is substantially smaller than the value used to compute Inequality 4.5, while the average value of the function has increased over the value used to compute Inequality 4.2. The bounds on the error due to integrating $A_1(x,t)$ are

$$\varepsilon_I \leq 8.1 \times 10^{-4},$$

$$\varepsilon_d \leq 10^{-4},$$

where the value of ε_d was computed under the optimistic assumption that the data to be integrated has three significance digits. The total bound on the error in the second integration of $x(t)$ is then given by

$$\varepsilon_I^2 \leq 9.1 \times 10^{-4}, \quad (4.10)$$

and this bound questions the accuracy of the third digit of the data which in notational form is denoted by $A_2(x,t)$.

For the third and higher orders of integration, the average value of the function being integrated remains essentially constant; however, according to the computed bounds $A_2(x,t)$ has at most two significant digits. Computing the bounds on the errors introduced when $A_2(x,t)$ is integrated to form $A_3(x,t)$ it is found that $A_3(x,t)$ has at most two significant digits while the fourth and higher orders of integrations have at most one significant digit.

Estimating the Error Bounds by Comparison

A more practical approach for determining the uncertainty in the results of integrating numerically is to compare the results of using different sampling intervals. That is, half and/or double the sampling interval and recalculate the integrated quantities. All results are then compared in order to estimate the number of significant figures. In practice, this method of checking the accuracy of the integration process will actually be easier to automate; however, both methods can be employed as a cross-check.

The impulse response of the system is generated at sampling

intervals of 0.01 and 0.005 and 0.0025 seconds, and the sampled signals integrated as before. Table 2 indicates the results of four orders of integration using all three step intervals. The values given in the table are results of integrating for two and one quarter seconds in each case. It is seen that the results for the step intervals of 0.005 and 0.0025 seconds agree to three digits in all cases. These results in turn agree with the results for a step interval of 0.01 seconds to at least two digits. There is thus reason to believe that the step interval of 0.005 seconds allows at least three significant digits in the integrated data.

Table 2. Comparison of Integrating with Different Step Intervals

Sampling Interval	$A_1(-x, 2.25)$	$A_2(-x, 2.25)$	$A_3(-x, 2.25)$	$A_4(-x, 2.25)$
0.01	0.024144	0.072577	0.091219	0.070371
0.005	0.024170	0.072678	0.091336	0.070456
0.0025	0.024176	0.072639	0.091301	0.070434

Solving a System of Simultaneous Equations

The generated table of data will now be used to establish systems of equations and various methods will be used to solve these equations. After first discussing the general way in which these equations can be solved, and then giving results of using these ways, the final selected method by which the required system of equations are to be solved in connection with the identification procedure will be given.

General Methods for Solving Systems of Equations

The available methods for solving a system of linear algebraic equations which are applicable to the problem of this research can be classified as

- (1) Direct methods,
- (2) Successive approximation methods.

Included in the direct method is Gauss elimination and the method of matrix inversion, while the successive approximation method includes the method of steepest descent and the conjugate gradient method.

There are several considerations to be given to the choice of the method selected for solving a system of equations. Discounting the situation in which it is necessary to obtain the best approximate solution with a minimum of computer storage and time, the most significant aspects of the problem are reducing roundoff errors, and converging to the true solution. Usually with direct methods, if it were not for roundoff errors, it would be possible to obtain an exact solution, for an exact system of equations, after a finite number of mathematical operations. With the exception of the fairly recently developed conjugate gradient method, most of the successive approximation methods will not yield, even theoretically, the exact solution in a finite number of steps.

While the above discussion is for the case of an exact system of equation, in practice it is only possible to approach the exact case and thus it is necessary to obtain the least squares solution. The method of solving the system of equations must then be chosen so as to yield this desired result.

Direct Methods. The direct methods considered for this research are

- (1) Crout reduction.
- (2) Matrix inversion.
- (3) Gauss elimination.

The Crout reduction and the matrix inversion methods are available as programmed procedures through Burroughs Corporation^{(29),(32)}. An iterative procedure⁽³³⁾ for reducing roundoff error in an inverted matrix is also available and is used in conjunction with the matrix inversion routine. The Crout reduction procedure employs pivoting and has a built in iterative scheme to reduce roundoff error in the solution. The Gauss elimination procedure is written in double precision and employs pivoting only when a zero element occurs on the diagonal. The procedure does not iterate upon the solution in order to reduce roundoff error; however, the use of double precision arithmetic helps to hold rounding error to a minimum.

All three direct methods have been employed in several tests in order to compare results, and the results of a typical test is given in Appendix D. In all tests, the same number of equations as unknowns are selected from the total set of available equations. The number of unknowns are consistent with the order of the system which the table of data represents, and the equations are solved directly without the formation of the normal equations*. The conclusions of testing with several problems and comparing results with the known true solutions

* Systems of equations of size $(T \times k)$ which are not converted into the normal equations will be denoted as regular systems of equations throughout the remainder of this work.

are that either of these direct methods can be successfully used in an implementation of the identification procedure. A direct method does not necessarily yield the least squares solution; however, if the system of equations are close to being exact then the solution obtained by a direct method will be close to the least squares solution.

Successive Approximation Methods. The approximation methods are discussed extensively in the literature, both from a theoretical aspect and from conclusions formed during actual usage of the procedures*. An iterative method is particularly desirable in this work because these methods in general minimize the residue vector during their iterations. Two iterative methods have been considered for this work. These are the conjugate gradient method as given by Bechman⁽³⁴⁾, and the method of steepest descent as given by Booth⁽³⁵⁾. The best method for solving a system of linear equations, according to several authors, still appears to be a straight Gaussian elimination method**. While the elimination method requires a more complex computer program than does either of the approximation methods, the approximation methods require in general, larger computer storage facilities.

Just as in the case of a direct method, roundoff error effect the solution obtained by the conjugate gradient method. The extent of rounding error in the method of conjugate gradient is related to the extent of the ill-conditioning of the system of equations; however, it is possible to restart the conjugate gradient method at any time and reduce the effect of error growth. Error growth is not serious in the

* Fischback(36), pp. 59-72; Hestenes (36), pp. 83-102; Rosenbloom (36), pp. 127-176; Young (36), pp. 283-298; Bechman (34).

** Fischback (36), pp. 59-71; Bechman (34).

method of steepest descent; however, unless the equations are well conditioned each new iteration tends to oscillate about the most direct path towards the solution.

In theory, the method of conjugate gradient will converge to the true solution for an exact system of equations in k iterations, where k is the order of the system of equations. In actual use, it has been found that roundoff error prevents the method from yielding the exact solution in k iterations and frequently the $(k + 1)^{st}$ iteration is better than the k^{th} iteration*. If the system of equations is extremely ill-conditioned, error growth can be so serious as to require an additional k iterations to correct the solution. Again, a system of equations established in connection with the identification procedure will not have an exact solution and more than k iterations will in general be required in order to obtain the least squares solution. It is of interest to note that the conjugate gradient method can theoretically yield the least squares solution even when the coefficient matrix is singular**.

Both of the iterative methods must be given starting values, and for an identification problem the only general choice of starting values are zeros. Each of the iterative methods have been tested in a manner similar to the tests conducted with the direct methods, and typical results are given in Appendix D. The results of these tests are:

1. The method of steepest descent was always slow in converging to a solution, and the obtained solutions were in general physically

* Bechman (34).

** Bechman (34).

meaningless when compared to the solutions obtained by a direct method, and to the known true solutions. The conclusions of all testing with the method of steepest descent is that it is entirely unsatisfactory for use in connection with the identification procedure.

2. The conjugate gradient method readily converged to a solution in all tests. The norm of the residue vector was always relatively small; however, in some cases the resulting solution was far from being correct. While the conjugate gradient method will sometimes yield the correct solution, because of its uncertainty it cannot be entirely accepted as a method for solving the identification problem. Solutions obtained when using the conjugate gradient method readily illustrates the results of Equation 3.10. That is, the norm of the residue vector would be relatively small; however, the solution would be far from the true least squares solution.

Obtaining the Least Squares Solution

As discussed previously, it is desirable to obtain the least squares solution for an established system of equations, and it is also desirable that the number of equations not be restricted by the number of unknowns. Thus, the system of equations can be of size $(T \times k)$, for $T \geq k$. In theory, the least squares solution can theoretically be obtained by forming and solving the normal equations.

The formation of the normal equations has been considered as a possible method for obtaining a solution to the identification problem; however, results of tests have been poor. Even when working with k equations in k unknowns, better results have always been obtained by solving the regular equations with a direct method than have been

obtained by solving the normal equations. Appendix D contains typical results of these tests. If the number of equations is increased, with the same number of unknowns, the solution obtained from solving the normal equations have been even poorer than results obtained when using only k equations in k unknowns.

The fact that the normal equations are not as reliable as the regular equations is explained by the conditioning of the two systems of equations. When two matrices are multiplied together, the conditioning of the product can be no better than the conditioning of the individual matrices and in general it will be worse. This fact will be illustrated in a later section when the conditioning number of several systems of equations are compared.

Both the conjugate gradient method and the method of steepest descent can be used to solve T equations in k unknowns; thus it seems reasonable to employ one of these methods if possible. If the two methods are examined, it is found that the normal equations are formed in the iterative process of the method of steepest descent; however, the normal equations are not explicitly formed in the method of the conjugate gradient.

The method of steepest descent has been rejected in connection with the identification procedure; however, experimentally it has been found that when given any reasonable starting values, the conjugate gradient method will reduce the norm of the residue vector, and improve the solution. Appendix D gives a typical result of a test in which the starting values are vague approximations to the true solution. The example illustrates the rapid convergence of the conjugate gradient

method. Thus, the least squares solution can now be obtained for a system of $(T \times k)$ equations if a reasonable good estimate of the solution is available.

If a preliminary solution is first obtained by solving a system of k equations in k unknowns by a direct method, it is then possible to use T equations in k unknowns, in conjunction with the conjugate gradient method to iterate upon the preliminary solution. The results will be the least squares solution for the total set of T equations. Appendix D contains an example illustrating this method.

The chosen method of solving the identification problem has two advantages. The first is that any desired number of equations can be used, and the second is that the use of a direct method will allow the computation of the uncertainty of the resulting solution. The method of determining the uncertainty of a solution, as given in Chapter III, relies upon the fact that the solution is obtained from a system of equations of size $(k \times k)$ with the use of a direct method. Since the solution obtained by a direct method is now iterated upon and changed, the value of the method for determining the uncertainty can be questioned; however, it has been found in practice (as illustrated in Appendix D) that the preliminary solution is not radically changed by the iteration process. Thus the final solution and the preliminary solution are essentially the same. This result is reasonable because the complete set of equations does have a solution which yields a relatively small residue vector, thus a relatively correct solution should be obtained when a subset of the total set of equations are solved by a direct method. The closer the entire system of equations are to being true, the truer will

be a solution obtained by solving a partial set of these equations.

In general, a solution obtained by solving k equations in k unknowns by a direct method has been very good. The necessity of iterating at all might then be questioned, and if facilities do not allow for iteration it can be omitted. However, examples have been found where the preliminary solution has been substantially improved by iterating. Also, since iterating tends to reduce roundoff error in the preliminary solution, it is desirable to iterate to some degree. The number of equations required for successfully solving the identification problem, as well as the method for selecting these equations will be discussed in the next chapter.

Uncertainty in the Results of the Identification Procedure

Chapter III presented a numerical method for estimating the uncertainty in the final solution of the identification procedure. However, just as in the case of estimating the bound on the total error introduced when a signal is integrated by a numerical method, the computed bound on the uncertainty of the final solution can also be pessimistic. An example of this pessimism will now be given, and as before a comparison method for determining the reliability of the final solution will be given.

Estimating the Uncertainty Using Numerical Methods

It has been estimated that a system of equations, established for the purpose of identifying the fourth order system of Table 1, will have approximately three significant digits. Table 3 gives the solution obtained when a system of equations were established and solved for the

test example. The table also lists the true solution which would be obtained if the established system of equations were exact, as well as the difference between the true solution and the determined solution.

Equations 3.14 and 3.15 have been used to compute the uncertainty in the calculated solution, and these values are also listed in Table 3. The calculations involve the elements of an inverse matrix as well as the elements of the computed solution. It is also necessary to specify the number of significant digits in the elements of the original equations. The computed uncertainties are listed in Table 3 with the number of significant digits "d" carried as an unknown.

According to previous estimates, the value of d is three; however, in examining Table 3 it is seen that for this value of d, all of the computed uncertainties except δc_1 are larger than the corresponding computed solutions. It is also seen that if d has a value of nine, then each computed uncertainty will have a minimum value and yet be larger than the corresponding true uncertainties. However, it is impossible to have nine significant digits in the data since over half of the elements were truncated to approximately five digits.

Further examination of the obtained solution shows that if the solution is truncated to three significant digits then the computed solution and the true solution are identical. Even if the computed solution is rounded to three digits, it is almost exact. Thus it seems that the original estimate of three significant digits has some basis.

As seen by this example, the computation of the uncertainty of a solution can yield pessimistic results. This pessimism is due to

Table 3. Uncertainty in the Solution of a System of Equations

	True Solution	Solution by Matrix Inversion	True Uncertainty	δc_i as Computed Using Equations 3.14 and 3.15
C_1	1	1.0000	0.0000	0.00045×10^{-d}
C_2	7	7.0089	0.0089	$0.05 \times 10^{9-d}$
C_3	16	16.027	0.027	$0.04 \times 10^{9-d}$
C_4	10	10.024	0.024	$0.096 \times 10^{9-d}$
C_5	19	19.009	0.009	$0.049 \times 10^{9-d}$
C_6	118	118.11	0.11	$0.75 \times 10^{9-d}$
C_7	320	320.38	0.38	$2.45 \times 10^{9-d}$
C_8	400	400.56	0.56	$3.78 \times 10^{9-d}$

the fact that the inverse matrix contains relatively large elements, and thus the resulting solution can have a large degree of uncertainty.

Again, in order to reduce the range of uncertainty, it is necessary to carry a large number of significant digits in the data. If this is impossible, there is still another method for checking the reliability of the obtained solution. This method is outlined next.

Estimating the Reliability by Comparison

Table 2 illustrated a "comparison" method for determining the accuracy of an integration method employed to numerically integrate a sampled signal as well as for determining the best sampling interval to be used. A similar comparison method can be used in connection with the problem of determining the reliability of the final solution of the identification problem. This comparison method is to solve the

identification problem using two or more different sampling intervals and/or integration schemes, and compare results. Table 4 lists two solutions obtained by solving the same identification problem using two different sampling intervals. The solution for $H = 0.005$ is the solution listed in Table 3, while the solution for $H = 0.01$ has been obtained by solving the identification problem using the appropriate table of data.

Table 4. Comparison of Results Using Two Different Sampling Intervals

	True Solution	Solution Obtained for $H = 0.005$ sec.	Solution Obtained for $H = 0.01$ sec.
C_1	1	1.0000	1.0000
C_2	7	7.0089	7.0068
C_3	16	16.027	16.051
C_4	10	10.024	10.057
C_5	19	19.009	19.017
C_6	118	118.11	118.16
C_7	320	320.38	320.49
C_8	400	400.56	400.66

If the data generated for a sampling interval of $H = 0.005$ has three significant digits then according to the comparison of Table 2 the data generated for $H = 0.01$ should have only two significant digits. Table 4 illustrates that both sampling intervals yields approximately the same results and that both results are true to three significant

digits. An ill-conditioned system of equations is defined as one in which small variations in the elements of the equations will produce large variations in the final solution. It is seen by this example that although the elements of the inverse coefficient matrix are large, the solution of the system of equations does not necessarily have to yield incorrect results because the elements of the equations are perturbed slightly.

Tables 2 and 4 both illustrate methods for determining the accuracy of a numerical integration scheme which is to be used in connection with the identification procedure. In Table 2 the required integration is performed using three different sampling intervals and the results are compared to determine the best interval. In Table 4 the identification problem is solved with data obtained by using two different sampling intervals before the results are compared. While it is true that the comparison of Table 2 can be an intermediate step for comparing the results of Table 4, it is also true that the final desired results is a correct solution to the identification problem. Thus, when the available data does not have enough significant digits to allow the uncertainty of the final solution to be determined with any reliability by using a numerical technique (as discussed in Chapter III), then other means must be employed. Table 4 illustrates one such method for estimating the reliability of a solution; however, other methods can be employed. This problem will be discussed more in the next chapter.

Determination of Order

A basic problem in solving the identification problem is to determine the order of the system. In the theoretical development of Chapter II, it is proven that the maximum rank of an established system of equations is directly related to the order of the system being identified, and the solution of the largest non-singular system of equations will then yield the parameters of the correct order model. In theory, the procedure would then be to first establish a system of equations for a model of arbitrary order s , and test the coefficient matrix to determine if it is singular. The largest non-singular system of equations which can be formed would then be solved for the parameters of the correct order model. However, in practice it has been found that the above procedure cannot be successfully implemented.

Even when the coefficient matrix is truly singular, because of roundoff error, it is actually possible in practice to sometimes solve this singular system of equations. This undesirable result is inherent in all numerical methods for solving algebraic equations. When solving a system of equations by elimination (or by matrix inversion), the coefficient matrix is first triangularized. During the triangularization process, one column of the triangular matrix will in theory contain all zero elements below the diagonal if the equations are dependent. However, computationally it is found that the column will not actually contain all zeros, but will contain very small element values. The actual test for singularity is then to compare the largest element, in absolute value, against a small test number. If all elements in a column are in absolute value less than the specified test number, then

the matrix is declared to be singular.

Although there appears to be no available method for determining the order of an unknown system, in practice this is not a problem. Just as in the theoretical proof where a system with surplus factors is a solution for a larger system of equations, in practice, the solution of the larger system of equations does have surplus factors and when these factors are removed the resulting solution will reduce to the correct model -- both with respect to order and parameter values.

Because of surplus factors, it is desirable to actually solve an established system of equations for models of the same order and of higher order than the suspected order of the system. When all surplus factors are removed, the higher order and same order attempts (as the true order of the system) will then reduce to the same essential model. With fairly good data (such as the data for this test system), this has been the results obtained in practice. This result will tend to give more confidence in the resulting model. This point will be illustrated in the next chapter with examples, and consideration will be given as to the required accuracy of the data necessary before the identification problem can be solved.

Surplus Factors

In theory, a surplus factor occurs when a pole and a zero of a system are identical; however, in computational work it will be found that a surplus factor must be defined as a pole and a zero which are arbitrarily close. It is necessary to define what is meant by closeness, and it will be said that a pole and a zero are surplus factors whenever

their removal will have a negligible effect upon the response of the model.

Assume that a signal of the form $y = |Y| \sin(\omega t)$ excites a given model. The gain of the model in decibels is then given by

$$G_1 = 20 \log_{10}(E_1/Y) , \quad (4.11)$$

where the response of the system has the form $e_1 = |E_1| \sin(\omega t + \Phi)$. If this model has what is believed to be a surplus factor then the removal of this factor will result in a gain of

$$G_2 = 20 \log_{10}(E_2/Y) , \quad (4.12)$$

and a ratio then exists of the form

$$\frac{E_1}{E_2} = 10^{[(G_1 - G_2)/20]} \quad (4.13)$$

If the absolute value of the difference between G_1 and G_2 is less than one decibel then

$$0.990 \leq \frac{E_1}{E_2} \leq 1.01 \quad (4.14)$$

Consider now a system which has only real poles and zeros in the left hand plane of the complex frequency plane. If the ratio of the "distance" between a pole and a zero to the absolute value of the real part of the pole or the real part of the zero (which ever is greater) is less than 0.1 then the removal of this pair of roots will effect the response of the system at most by one decibel at any frequency. If

both the pole and the zero have real parts which are in absolute value less than one, then the criterion is to just examine the distance between the two roots. The value of 0.1 as a test of closeness is an arbitrary choice, and it can be either reduced or enlarged. Also, before any arbitrarily close pole and zero can be removed, their effect upon the response of the model must be considered with respect to the other poles and zeros.

For the work of this research, a ratio of 0.05 has generally been sufficient to remove surplus factors, and this value has been used in an automated method for removing surplus factors. For a ratio of 0.05, the magnitude of the frequency response of a model will change at most by one-half a decibel at any frequency. The bound on the ratio of the output signal at any frequency then becomes

$$0.9970 \leq \frac{E_1}{E_2} \leq 1.003 . \quad (4.15)$$

Although the above discussion is for the case of a stable model with real poles and zeros, the same criterion for removing surplus factors can also be applied to models which have complex roots as well as roots on the imaginary axis and in the right hand plane of the complex frequency plane. If a model should have an identical pole and zero located either on the imaginary axis or in the right hand plane then the response of this model will be bounded for all time t . Since surplus factors are not restricted to lie only in the left hand plane, and because of computational errors which prevent surplus factors from always being identical it is then necessary to use the above criterion at all times when checking for surplus factors. In fact every effort should

be made to remove roots which cause the response of the model to become unbounded even to the point of relaxing the test requirements. This is done in order to produce a stable model.

The Condition of Different Systems of Equations

Appendix D contains the results of solving a regular system of equations as well as the results obtained when the normal equations are formed and solved. In Appendix D it was found that the normal equations did not yield in general as good results as the regular equations, and the reason given is that the "condition" of the normal equations is worse than the "condition" of the regular equations. This section compares the conditioning number of several systems of equations which have been formed and solved in connection with the identification procedure and compares the results to theory.

The spectral conditioning number, as defined by Equation C.34 in Appendix C is

$$\eta = \|A\|_3 \|A^{-1}\|_3, \quad (4.16)$$

where $\|\cdot\|_3$ is the matrix norm defined by Equation C.11. Computationally, it is easier to compute the Euclidean norm which is denoted by $\|\cdot\|_E$, and which is defined by Equation C.12. The Euclidean norm will be used in this discussion; however, from Inequality C.13 it is possible to write the following inequality

$$\frac{1}{k^{1/2}} \|A\|_E \leq \|A\|_3 \leq \|A\|_E \quad (4.17)$$

where k is the order of the matrix A . The spectral conditioning number is then bounded by

$$\frac{1}{k} \|A\| \|A^{-1}\|_E \leq \eta \leq \|A\|_E \|A^{-1}\|_E. \quad (4.18)$$

For the fourth order example used in Appendix D, when eight equations are established in eight unknowns, the conditioning number is bounded by

$$3.6 \times 10^6 \leq \eta \leq 28.5 \times 10^6. \quad (4.19)$$

If the equations are transformed into the normal equations, the conditioning number is larger, and is bounded by

$$3.4 \times 10^{13} \leq \eta \leq 27 \times 10^{13}, \quad (4.20)$$

These conditioning numbers are for the system of equations in Appendix D, where the regular equations yielded essentially the correct solution, while the normal equations did not yield as accurate a solution. It is seen that the conditioning number of the normal equations is several orders of magnitude larger than the conditioning number of the regular equations. Thus theoretical and experimental results agree.

If an attempt is made at finding a fifth order model the conditioning number of the tenth order matrix is bounded by

$$3.67 \times 10^{11} \leq \eta \leq 36.7 \times 10^{11}. \quad (4.21)$$

This conditioning number is larger than that of the eighth order regular equations, but it is not as large as the conditioning number of the

eighth order normal equations. The tenth order system of equations, in theory, should be singular. The fact that the tenth order system of equations does have a larger conditioning number than the eighth order system of equations does indicate that the larger system of equations is more likely to be singular. The fact that the tenth order regular equations yielded better results than the eighth order normal equations also agrees with the fact that η of Inequality 4.21 is smaller than η of Inequality 4.20.

If the normal equations are formed from the regular tenth order equations, its conditioning number is bounded by

$$4.75 \times 10^{14} \leq \eta \leq 47.5 \times 10^{14} \quad (4.22)$$

This is the largest η of all cases, and again this agrees with theory because the solution of these equations were meaningless.

CHAPTER V

IDENTIFICATION: GIVEN IMPULSE RESPONSE

This chapter includes an investigation of the identification procedure for the case where the system is excited by an unit impulse. Use is made of the automated identification procedure in order to gather data for examples. Further discussions are also included on certain significant aspects of the identification procedure.

Response Length Required for Identification

The theoretical development of the identification procedure placed no bounds on the time length of the response which must be available for successful identification. It is now necessary to investigate this aspect of the identification procedure in more detail.

To be specific, assume that an impulse response is measured for T seconds, where T is some value of time approximately equal to the essential time length of the total response. The procedure developed in the previous chapter is to first select k equations and then solve for k unknowns by a direct method. Here it will be assumed also that k is twice the order of the unknown system and that these equations can be solved for the correct parameters of the system.

Now, the value of T is decreased in steps of ΔT , and identification is attempted for each new value of T . Since the number of equations must remain fixed, it is obvious that the "distance" between each equation is being diminished and the equations are becoming more and more "parallel."

The practical consequences are that the equations become dependent in the limit as T decreases to zero, and it is thus impossible to numerically solve the system of equations for T less than some value T' .

Thus, while the theoretical development placed no bounds on the value of T' , physical consideration does limit the minimum time length of the available response if identification is to be attempted. In practice, physical considerations not only limit the required available time length of the response to something other than a vanishing small interval, they require that a large percentage of the total response be available. This is explained as follows.

In general, the poles of a system can be separated into groups according to the value of their real part, and each grouping will tend to dominate some portion of the total response. For example, the smaller roots will tend to determine the essential time length of the total response, while the larger roots will tend to dominate the early portion of the response. Thus, if only the early portions of the total response is available, it is very likely that the effect of the smaller roots can be approximated by a fewer number of roots.

Experimentally it has been found that the results of solving a system of equations is that the solution will tend to reduce to the smallest order model, and yet still closely approximate the response of the system. Even when attempting identification for the correct order model, if the established system of equations are not truly representative*

* By representative is meant that the equations are selected over the entire essential time length of the response rather than from a smaller portion of the response. The available response time should not be confused with the essential response time. The available response time is defined for this work as the time length of the response which has been measured in order to obtain data for the identification procedure. For successful identification it is generally necessary that the response be measured for the total essential time length.

then the solution of the system of equations will contain surplus factors and reduce to a lower order approximate model. To overcome this problem, it is necessary to establish a system of equations which are truly representative, in order to allow all the roots of the system to have an influence upon the final solution. This point will be illustrated with examples.

Establishing the System of Equations

Both the number of equations required for the identification procedure and the method of selecting these equations have been investigated in great detail. As discussed above, the preliminary system of equations of size $(k \times k)$ must be selected from the total essential response length rather than from segment of the total essential length. The easiest method is to select the equations on a linear time basis; although, any selection which approximates a linear time basis is satisfactory. It has been found experimentally that the preliminary equations usually yield excellent results, and thus only a minimum number of iterations are then required.

Iterating upon the preliminary solution is recommended even when the equations used for iterating are simply the preliminary set of equations. If more equations are to be used in the iterating process, it is recommended that these equations be selected on a quantized basis. That is, the response is quantized into equal levels, and an equation is selected each time the signal has a value equal to one of the quantized levels. In this way, more equations can be selected during the portions of time where the signal exhibits large variations.

There are no restrictions as to how many equations must be used

in conjunction with the conjugate gradient method for the purpose of iterating, and in practice it has been found that it is not necessary to use a large number of equations. As an example, for the fourth order system considered in the previous chapter, just as good results have been obtained when using approximately 25 equations as when using 200 equations. A good rule is to select approximately three or four times as many equations as unknowns.

Gradient of the Error Function

The error function defined as ERMS by Equation 2.2 will be zero only when the model is a perfect representation of the system. In practice it will generally be impossible for ERMS to be zero because the measured response of the system will be noisy. It is then desirable to vary the parameters of a model so as to minimize the response error.

If the response of a model is computed for a set of parameters $\hat{\underline{c}}$, then the response of this model can be written in notational form as $z(\hat{\underline{c}}, t)$. The effect of varying the parameters $\hat{\underline{c}}$ is to vary the value of the error function. This can be written as

$$\partial \text{ERMS} / \partial \hat{c}_i = F[x(t), z(\hat{\underline{c}}, t), \partial z(\hat{\underline{c}}, t) / \partial \hat{c}_i] , \quad (5.1)$$

where $F[\cdot]$ is a function of both the response of the system, and the response of the model. The set of $\left\{ \partial \text{ERMS} / \partial \hat{c}_i \right\}$ for all \hat{c}_i form the gradient of the error function, and is generally required if the minimization is performed in conjunction with a multivariable search procedure.

An alternate method for solving an identification problem is to employ multivariable search procedures, a discussion of these procedures along with their possible use in conjunction with this research is included in Appendix E. The one problem with employing multivariable

search procedures is that the order of the model along with initial estimates of the parameters of the model must be known. The identification procedure of this research readily yields both a correct order model and its parameter values. Thus, when the measured data is such that the parameters have not been determined with reasonable accuracy, multivariable search can then be employed to improve the model. Multivariable search techniques and the identification procedure of this research thus complement each other.

Automation of the Identification Procedure

A computer program has been written for the B5500 computer in order to be able to systematically investigate the identification procedure. The program was developed concurrently with the implementation and investigation of the identification procedure. Although the program does not represent a completely automated version of the identification procedure, it does represent the basic way in which the procedure must be automated. Specifically, the program is written only for the aspects of establishing and solving the required system of equations once the proper data is available, and very little automation has been included for the determination of the uncertainty of a resulting solution. Whenever the uncertainty of the integrated data and its effect upon the solution has been investigated, it has been done by writing special programs or by making use of parts of the basic identification program.

Figure 4 is a block diagram representation of the basic computer program including the steps used to generate the test data as well as

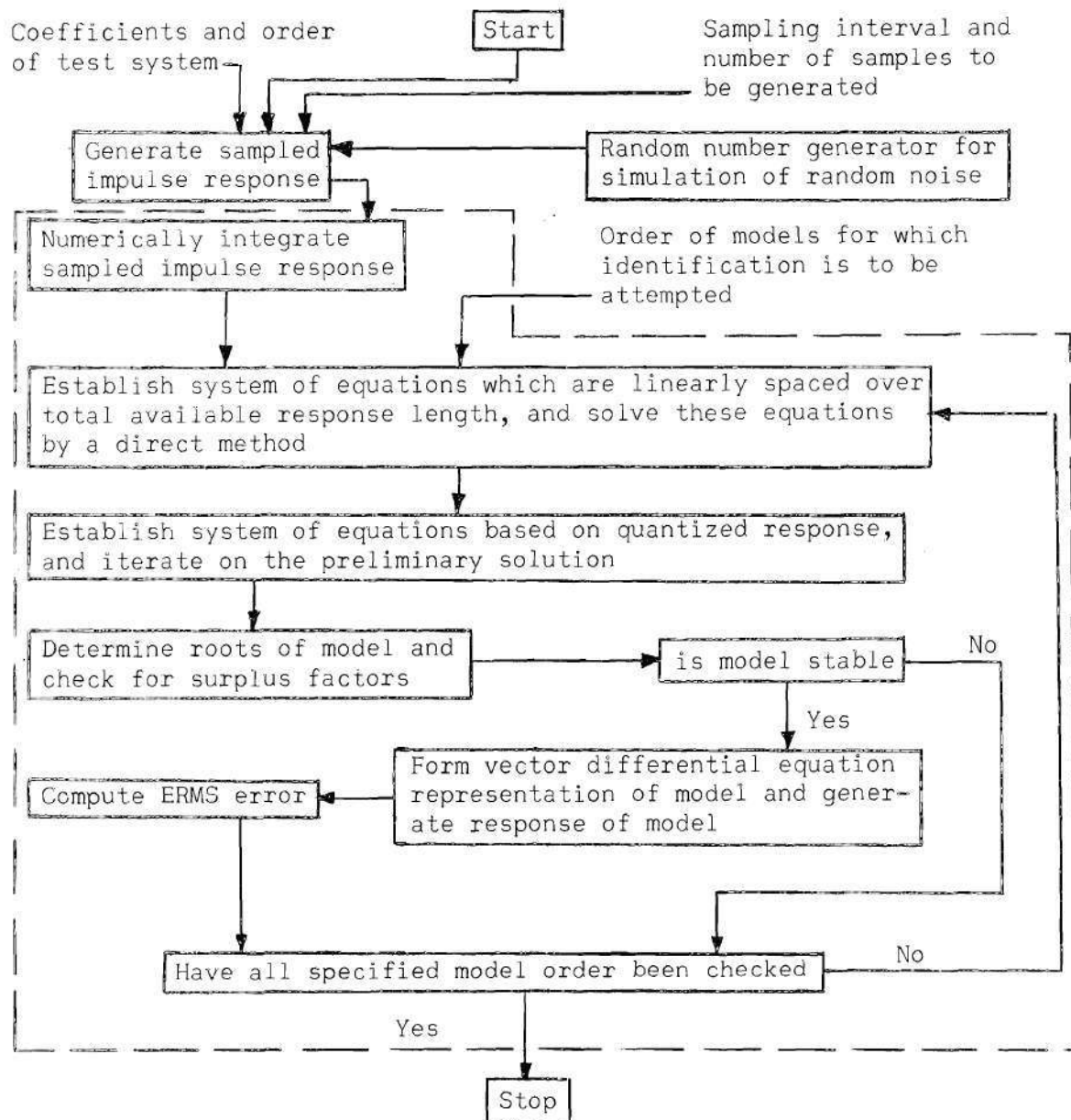


Figure 4. Block Diagram of Basic Identification Procedure and Input Simulator.

the steps in the process of solving an identification problem. The actual automated identification program is enclosed in dashed lines. The input to the data generator includes the order of the test system, the value of its coefficients, the sampling interval to be used, and the number of samples which are to be generated. Provisions are also included whereby a random number generator can be used to randomly disturb each sample value whenever test data are generated within the program.

The sample values, the sampling interval, and the number of samples are the input data to the identification program. Also included as input data is a specification of the order of the models for which identification is to be attempted. The program then establishes the required systems of equations, and solves for every specified order model. Checks are made for surplus factors, and the response of each stable model is generated. The ERMS error is then calculated for each model. In the actual program, provisions are included for printing the results of each step on the line printer. This then allows a visual inspection of all results.

Examples

The remainder of this chapter is concerned with presenting experimental results of an investigation of the identification procedure for the case when the impulse response of a system is available for measurement. Use is made of the program shown in Figure 4. The investigation is conducted using data obtained from several simulated systems; however, only the results of the investigation using the system of Table 1 will be discussed. The results are typical, and provide an insight

into the practical problems encountered in identifying a system. The examples are divided into two sections. The first section will discuss attempted identification using the generated data as discussed in Chapter IV. In the second section, the generated response is disturbed so as to simulate the effect of additive noise.

The automated identification procedure is presented in various time lengths of the response of the test signal, and in each case the program is directed to determine a fourth, a fifth, and a sixth order model. In each case, the proper number of equations are selected on a linear time basis and solved by matrix inversion. Approximately twenty-five equations are then selected on a quantized basis and used in conjunction with the conjugate gradient method to iterate on the preliminary solution. The roots of the final solution are determined, and these roots are then rounded to two places past the decimal.

The determined roots of the model are then checked for surplus factors using a maximum ratio of 0.05 if both the pole and the zero lie in the left half of the complex frequency plane. If either a pole or a zero lies in the right hand plane, then the test ratio is set at 0.1. After all surplus factors have been removed, the coefficients of the resulting model are then formed and the Runge-Kutta integration procedure is used to generate the response of the model. The response of the model is generated at every sample time for which the measured response of the system is available. The ERMS error and the performance index are then computed using Equations 2.4 and 2.5.

Examples Using Essentially True Data

The data for the system of Table 1 has been estimated as having

three significant digits, and for the investigation this data will be said to be relatively accurate. Results of identification attempts using this data can be divided into two classifications -- that is, successful results and unsuccessful results. The next two subsections will discuss these results.

Successful Identification. Identification attempts when the response is available for time lengths of (0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25) seconds can be classified as successful. In every identification attempt, the result is a fourth order model with approximately correct coefficient values. That is, both the fifth and the sixth order attempts had surplus factors and reduced to a fourth order model, and all fourth order attempts also yielded a fourth order model.

Table 5 lists the ERMS error and the performance index for all attempts. The individual coefficients of each model have been compared to the respective coefficients of the test system on a basis of absolute percentage deviation. Table 4 then lists the largest absolute percentage deviation of all the individual coefficients as well as the average absolute percentage deviations of all computed coefficients for a model.

A general statement of accuracy is that the results of each order of attempted identification is truer, with respect to coefficient value, for the larger time lengths of available response. This result emphasizes the statement made earlier concerning the required time length of available response for successful identification, and this point will be illustrated throughout the presentation of the examples.

It is also seen by inspecting the results of Table 4 that, for a specific time length of available response, the model which is truer

Table 5. Comparison of Results of Identification Attempts Using Essentially True Data

(a). Sixth Order Attempt

Time length of Available Response (seconds)	Sixth Order Attempt			
	<u>Per Cent Deviation</u>		ERMS	I
	Max.	Ave.		
2.25	2	1.1	2.8×10^{-3}	2×10^{-2}
2.0	1	0.4	1.4×10^{-3}	9.5×10^{-3}
1.75	3	1.5	4.7×10^{-3}	3×10^{-2}
1.50	10	4.8	1.7×10^{-2}	0.1
1.25	4.4	2.2	7.4×10^{-3}	3.9×10^{-2}
1.0	0.44	0.17	6.6×10^{-5}	3.1×10^{-4}
0.75	4.8	2	7.2×10^{-3}	2.9×10^{-2}

(b). Fifth Order Attempt

Time length of Available Response (seconds)	Sixth Order Attempt			
	<u>Per Cent Deviation</u>		ERMS	I
	Max.	Ave.		
2.25	0.2	0.085	1.3×10^{-4}	9.4×10^{-3}
2.0	0.2	0.15	1.7×10^{-4}	1.16×10^{-3}
1.75	0	0	0	0
1.5	1.3	0.42	2.3×10^{-4}	1.3×10^{-3}
1.25	0.3	0.26	1.7×10^{-4}	9.15×10^{-3}
1.0	0.44	0.16	2.3×10^{-4}	1.11×10^{-3}
0.75	1.6	0.8	2.7×10^{-4}	1.3×10^{-3}

Table 5. Comparison of Results of Identification Attempts
Using Essentially True Data (continued)

(c). Fourth Order Attempt

Time length of Available Response (seconds)	Sixth Order Attempt			
	Per Cent Deviation		ERMS	I
	Max.	Ave.		
2.25	0.4	0.2	1.6×10^{-4}	1.16×10^{-3}
2.0	0	0	0	0
1.75	0	0	0	0
1.5	0.43	0.37	1.2×10^{-4}	7.1×10^{-3}
1.25	0.3	0.13	1.9×10^{-5}	1.02×10^{-4}
1.0	3.0	1.05	7.7×10^{-5}	3.7×10^{-4}
0.75	4.0	1.7	2×10^{-4}	9.7×10^{-3}

with respect to coefficient values does not necessarily have the minimum ERMS error. It has been found that the ERMS functional, as defined by Equation 5.1, is very sensitive to small changes in the coefficients whenever all coefficients are very close to being true in value. On the other hand, whenever all coefficients are not as close to their true value, then the coefficient values can be more easily arranged so as to compensate for the total inaccuracies of all coefficients. Thus, the total ERMS error does not necessarily increase as all coefficients vary further from their true value. This aspect is particularly illustrated by the results of all attempts at identification when the response is available for a time length of one second.

The final models resulting from the fourth, the fifth, and the sixth order attempts for an available response time of one second are respectively

$$\begin{matrix} (4) & & (3) & & (2) & & (1) \\ x & + & 19.21 x & + & 120.29 x & + & 320.79 x & + & 395.73 x = \end{matrix} \quad (5.2)$$

$$\begin{matrix} (3) & & (2) & & (1) \\ y & + & 7.21 y & + & 15.74 y & + & 10 y \end{matrix} ,$$

$$\begin{matrix} (4) & & (3) & & (2) & & (1) \\ x & + & 19 x & + & 118.32 x & + & 320.19 x & + & 399.97 x = \end{matrix} \quad (5.3)$$

$$\begin{matrix} (3) & & (2) \\ y & + & 7.02 y & + & 15.93 y & + & 10y \end{matrix} ,$$

$$\begin{matrix} (4) & & (3) & & (2) & & (1) \\ x & + & 19.02 x & + & 118.18 x & + & 319.71 x & + & 399.18 x = \end{matrix} \quad (5.4)$$

$$\begin{matrix} (3) & & (2) & & (1) \\ y & + & 7.02 y & + & 15.93 y & + & 10.006 y \end{matrix} .$$

The average absolute percentage deviations of the coefficients of these models with respect to the true coefficients of the system are respectively (1.05, 0.16, 0.17), and the respective ERMS errors are (7.7×10^{-5} , 2.3×10^{-4} , 6.6×10^{-5}). It is seen that the fourth order attempt resulted in the worst model in terms of the closeness of the coefficient values; however, the ERMS error of this model is an order of magnitude lower than the error of the model resulting from the fifth order attempt. On the other hand, the sixth order attempt and the fifth order attempt produced models which have approximately the same average absolute percentage deviation in coefficient values, however, the model resulting from the sixth order attempt has the smallest ERMS error of all three models.

The surplus factors in all cases of fifth order attempts were far removed from the true root locations of the model, and in general the surplus factors were in both the right and left hand plane of the complex frequency plane. As an example of the size of surplus factors, the fifth order attempt for an available response time of one second had the following roots:

<u>poles</u>	<u>zeros</u>	
58,662*	58,698*	
$-1.99 \pm j1.99$	$-3 \pm j0.9$	(5.5)
-5.05	-1.02	
-10		
constant multiplier = 1.		

The surplus factors are indicated with an asterisk. It is seen that the remaining roots are very close to the roots of the test system as given in Table 1. It is also seen that the surplus factors are almost identical in value. Because of this, the remaining roots do not have to be shifted to compensate for the removal of these surplus roots. This closeness of all surplus factors was true for all fifth order attempts of Table 5 (b).

The sixth order attempts all had two sets of surplus factors, and in general one set was far removed from the true roots of the test system while the second set was close to the roots of the system. As an example, the solution of the sixth order attempt for an available response time of two and one-quarter seconds is

<u>poles</u>	<u>zeros</u>	
8.837*	8.838*	(5.6)
35,545*	34,895*	
$-2 \pm j2$	$-3 \pm j1.01$	
-4.99	-1	
-10		

Constant multiplier = 1.02

The surplus factors are indicated with an asterisk, and it is seen that the first set of surplus factors are close to being identical while the second set are not as close in value. If the second set of surplus factors were also identical then the remaining left hand plane roots as well as the constant multiplier must be changed in value in order that this solution will still be a least squares solution. In order to compensate for the removal of surplus factors which are not identical it is necessary to reiterate upon the parameters of the model formed after the surplus factors are removed.

To illustrate the above point, the parameters of the fourth order models resulting from the sixth order attempts of this example have been used as starting values for the conjugate gradient method. After eight iterations, the resulting model is then compared to the test system as before. Table 6 records these results. Comparing Table 6 to Table 5 it is seen that in every case the model resulting from the sixth order attempt has been improved both with respect to parameter values and ERMS error. The above form of iterating is recommended whenever surplus factors have been removed from the solution of an attempt at identification, although the additional iterating

is not necessary for the case where the surplus factors are identical.

Table 6. Results of Iterating on a Solution After Surplus Factors Have Been Removed.

Time Length of Available Response	Sixth Order Attempt		
	Per Cent Deviation		ERMS
	Max.	Ave.	
2.25	0.2	0.065	1.3×10^{-4}
2.0	0.2	0.065	1.4×10^{-4}
1.75	0.43	0.19	6.2×10^{-5}
1.5	0.44	0.26	2×10^{-4}
1.25	0.75	0.48	2.8×10^{-5}
1.0	0.4	0.17	7.9×10^{-5}
0.75	3.2	1.5	4.7×10^{-5}

Unsuccessful Identification. Identification attempts when the response of the system is available for lengths of (0.25, 0.5) seconds are classified as unsuccessful. The reason for this classification is that in all attempts, all models (except one) reduced to an approximation of a second order model. A second order model was also obtained whenever identification was attempted with a third order model for each of the indicated response lengths.

The best second order model with respect to the minimum ERMS error, is given by the following poles and zeros values:

$$\begin{array}{cc}
 \text{poles} & \text{zeros} \\
 -10.2 & -0.25 \\
 -2.04 &
 \end{array} \quad (5.7)$$

Constant multiplier = 1.

The impulse response of this model is given by

$$x(t) = 1.219 e^{-10.2t} - 0.219 e^{-2.04t}, \quad (5.8)$$

and it is seen that the two complex poles and the smaller real pole of the test system (as given in Table 1) have been approximated by a single real pole in this model. The pole of the test system which has a value of ten is the dominant root during most of the first one-half second of the response of the system, and as seen by Equation 5.8, this root also dominates the response of the model for this time interval. Again, all except one attempt reduced to a model very close to the model given above. Table 7 gives the ERMS error of each solution.

Table 7. Comparison of Results for Unsuccessful Attempts

Time Length of Available Response	Sixth Order Attempt ERMS	Fifth Order Attempt ERMS	Fourth Order Attempt ERMS	Third Order Attempt ERMS
0.5	2.8×10^{-3}	8.3×10^{-4}	1×10^{-4}	2.3×10^{-4}
0.25		3.9×10^{-3}	2.7×10^{-4}	2.5×10^{-3}

The one case where a second order model was not found occurred when identification was attempted for a fourth order model for the

available response time of one-half second. The final solution is given by

<u>poles</u>	<u>zeros</u>	(5.9)
878.8*	878.8*	
-9.07	-8.4	
-9.63	-0.15	
-1.84		

Constant multiplier = 1.

The surplus factors which were removed are indicated with an asterisk; however, it will be noted that another set of surplus factors exists if the test ratio is set at 0.1 rather than at 0.05. These factors are the pole at -9.07, and the zero at -8.4. When these roots are removed, the remaining roots are approximately equal to the roots of the second order model of (5.7). However, it should be noted that the response error of the third order model of (5.9) is smaller than the response errors of all results listed in Table 7.

Solving for a Third Order Model

The identification procedure is also applicable for the problem of determining a model which is of lower order than the true order of the system, but which will still closely approximate the response of the system. The degree of closeness will depend upon the system being approximated, the difference in orders between the system and the model, and the accuracy of the available data.

The identification procedure has been used to determine the parameters of a third order model for the fourth order test system.

As before, the required equations are first selected on a linear time basis from the total set of available equations, and solved by matrix inversion. Approximately 35 equations were then used for iterating. The iterating process did not change the first three digits of all coefficients determined by matrix inversion thus indicating that the preliminary solution is approximately correct.

The final solution is given by

$$\begin{matrix} (3) & & (2) & & (1) & & (2) & & (1) \\ x & + & 14.74 x & + & 50.182 x & + & 71.312 x & = & y & + & 2.62 y & + & 1.784 y, \end{matrix} \quad (5.10)$$

and the roots of the model are

$$\begin{array}{cc} \underline{\text{poles}} & \underline{\text{zeros}} \\ -2.04 \pm j1.59 & -1.31 \pm j0.26 \\ -10.66 & \end{array} \quad (5.11)$$

Constant multiplier = 1.

The ERMS error of this model as computed for the total time length of two and one-quarter seconds is

$$\text{ERMS} = 5.57 \times 10^{-4} \quad (5.12)$$

It is seen that this error is larger than the error of all fourth order models of Table 5, which were determined for the same time length of response; however, the error is relatively small when compared to the errors of Table 5.

Examples Using Noisy Data

A random number generator has been used to simulate the effect of additive noise on the response of the test system. Identification

is then attempted as before, and obtained results are compared to the known true solution. The random numbers are normalized so that they will be limited to lie within the interval $\pm B$, where the value of B is to be specified. The i^{th} random number is then selected and added to unity, and the i^{th} sampled is then multiplied by this number. In notational form, the process is

$$\overline{DP}_i \leftarrow DP_i (1 + rn_i), \quad (5.13)$$

where $DP_i = i^{\text{th}}$ sample of the true response of the system,
 $\overline{DP}_i = i^{\text{th}}$ sample of the noisy response,
 $rn_i = i^{\text{th}}$ random number.

It should be noted that this method of contaminating the signal does not correspond to an additive noise of constant variance since large values of the response can be disturbed the same percentage amount as small values of the response.

The RMS value as well as the average value of the additive noise is computed in order to obtain an indication of the size of the disturbance. The value of the noise at the i^{th} sample is given by

$$nv_i = DP_i - \overline{DP}_i, \quad (5.14)$$

and the RMS value is given by

$$RMS \Big|_{\text{noise}} = \left\{ \frac{1}{T} \sum_{i=1}^T |nv_i|^2 \right\}^{1/2} \quad (5.15)$$

where T represents the number of samples. Similarly the average value

of the additive noise is given by

$$\text{Average} \Big|_{\text{noise}} = \frac{1}{T} \sum_{i=1}^T nv_i. \quad (5.16)$$

The tests using the noisy response were conducted using three different values of B . In each case, the true response is first contaminated and then integrated according to the Trapezoidal rule. It is not expected that the identification procedure will yield an exact model for all attempts; however, it is hoped that the results will be such that the correct order of the unknown system will be determined along with good approximations to the value of the parameters. In this way, the preliminary model can be used in conjunction with a multivariable search procedure in order to improve the model. It should be noted that if an identification attempt should yield an exact model with respect to parameter values, then the minimum value of the ERMS error will be the RMS value of the additive noise.

Case No. 1: $B = 0.0025$. Table 8 gives an indication of the amount that the individual samples are disturbed when using a normalising factor of $B = 0.0025$. It is seen that anywhere from one to four digits of the original samples are unchanged, and in general two or three digits remain unchanged. Identification has been attempted when the response is available for time lengths of (2.25, 2.0, 1.75) seconds, and Table 9 lists the RMS value of the true signal as well as the RMS and average value of the additive noise for each time length of available response.

Table 10 summarizes the results of each identification attempt. The results in each case is a fourth order model, and each model is compared, with respect to coefficient values, to the test system. It is seen that the fifth order attempt yielded in each case a truer model than the respective fourth or sixth order attempt, although the fourth order attempt resulted in a model which has a smaller ERMS error. It is also seen that the sixth order attempt resulted in a model which in general lies somewhere between the respective fourth and fifth order attempts, with respect to coefficient values and ERMS error.

Table 8. Effects of Additive Noise for $B = 0.0025$

Sample Time (seconds)	DP_i	\overline{DP}_i
0.005	0.94154	0.94277
0.010	0.88609	0.88604
0.015	0.83348	0.83327
0.020	0.78356	0.78505
0.025	0.73662	0.73732
2.225	1.9317×10^{-3}	1.9336×10^{-3}
2.230	1.9158×10^{-3}	1.9170×10^{-3}
2.235	1.8997×10^{-3}	1.9025×10^{-3}
2.240	1.8836×10^{-3}	1.8824×10^{-3}
2.240	1.8674×10^{-3}	1.8687×10^{-3}
2.250	1.8512×10^{-3}	1.8521×10^{-3}

Table 9. Comparison of Uncontaminated Signal and Additive Noise for $B = 0.0025$

Time Length of Available Response (seconds)	RMS Value of Uncontaminated Signal	RMS Value of Additive Noise	Average Value of Additive Noise
2.25	0.14	1.7×10^{-4}	9.8×10^{-6}
2.0	0.15	1.75×10^{-4}	1.05×10^{-6}
1.75	0.16	1.87×10^{-4}	1.26×10^{-6}

Case No. 2: $B = 0.005$. Table 11 gives an indication of how the individual samples are disturbed for a value of $B = 0.005$. In general, at least two or three digits are left unchanged at each sample point. Table 12 lists the RMS value of the uncontaminated signal as well as the RMS value and average value of the additive noise.

Table 13 summarizes the results of each identification attempt. All attempts except the fourth order attempts for the available response times of (2.0, 1.75) seconds resulted in a fourth order model. A comparison of the coefficients of the fourth order models with the true system is given in Table 13. The two fourth order attempts for the above mentioned time lengths resulted in a third order model, and these models are compared to the third order model of Equation 5.10 with respect to coefficient values. These comparisons are indicated with an asterisk in Table 13.

Table 10. Comparison of Results for $B = 0.0025$

(a). Sixth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Sixth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	1.7×10^{-4}	1.8×10^{-4}	11	7.8
2.0	1.75×10^{-4}	4.2×10^{-4}	20	14
1.75	1.87×10^{-4}	2.4×10^{-4}	7.6	5.3

(b). Fifth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Fifth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	1.7×10^{-4}	2.1×10^{-4}	7.1	4.8
2.0	1.75×10^{-4}	4.5×10^{-4}	2.6	1.5
1.75	1.87×10^{-4}	2.9×10^{-4}	6.3	2.9

(c). Fourth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Fourth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	1.7×10^{-4}	1.9×10^{-4}	11	7.4
2.0	1.75×10^{-4}	2.9×10^{-4}	8.4	5.8
1.75	1.87×10^{-4}	2.2×10^{-4}	9.6	7.0

Table 11. Effects of Additive Noise for $B = 0.005$

Sample Time (seconds)	DP_i	\overline{DP}_i
0.005	0.94154	0.94399
0.010	0.88609	0.88600
0.015	0.83348	0.83306
0.020	0.78356	0.78654
0.025	0.73622	0.73843
2.225	1.9317×10^{-3}	1.9354×10^{-3}
2.230	1.9158×10^{-3}	1.9182×10^{-3}
2.235	1.8997×10^{-3}	1.9054×10^{-3}
2.240	1.8836×10^{-3}	1.8813×10^{-3}
2.245	1.8674×10^{-3}	1.8700×10^{-3}
2.250	1.8512×10^{-3}	1.8530×10^{-3}

Table 12. Comparison of Uncontaminated Signal and Additive Noise for $B = 0.005$

Time Length of Available Response (seconds)	RMS Value of Uncontaminated Signal	RMS Value of Additive Noise	Average Value of Additive Noise
2.25	0.14	3.3×10^{-4}	2×10^{-5}
2.0	0.15	3.5×10^{-4}	2.2×10^{-5}
1.75	0.16	3.74×10^{-4}	2.5×10^{-5}

Table 13. Comparison of Results for $B = 0.005$

(a). Sixth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Sixth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
3.25	3.3×10^{-4}	4.22×10^{-4}	28	20
2.0	3.5×10^{-4}	7.7×10^{-4}	35	24
1.75	3.74×10^{-4}	4.4×10^{-4}	5	2.3

(b). Fifth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Fifth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	3.3×10^{-4}	4.8×10^{-4}	16	11
2.0	3.5×10^{-4}	1.1×10^{-3}	57	38
1.75	3.74×10^{-4}	6×10^{-4}	12	7.2

(c). Fourth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Fourth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	3.3×10^{-4}	4.3×10^{-4}	19	13
2.0	3.5×10^{-4}	5.1×10^{-4}	22*	12*
1.75	3.74×10^{-4}	2.6×10^{-3}	18*	10*

Case No. 3: $B = 0.025$. Table 14 gives an indication of the amount each sample value is disturbed for a value of $B = 0.025$, and in general only one or two digits are left unchanged. Table 15 lists the comparison between the true signal and the RMS and average value of the additive noise.

Table 16 summarizes the results of each identification attempt. Most attempts resulted in a third order model, and the coefficients of these models are compared to the third order model of Equation 5.10. All of these comparisons are indicated with an asterisk. The remainder of the attempts resulted in a fourth order model, and these model are compared to the true system. It is noted that each attempt resulted in a model which has an ERMS error which is very close to the RMS value of the additive noise.

Table 14. Effects of Additive Noise for $B = 0.025$

Sample Time (seconds)	DP_i	\overline{DP}_i
0.005	0.94154	0.95378
0.010	0.88609	0.88565
0.015	0.83348	0.83139
0.020	0.78356	0.79845
0.025	0.73622	0.74726
2.225	1.9317×10^{-3}	1.9501×10^{-3}
2.230	1.9158×10^{-3}	1.9282×10^{-3}
2.235	1.8997×10^{-3}	1.9282×10^{-3}
2.240	1.8836×10^{-3}	1.8733×10^{-3}
2.245	1.8674×10^{-3}	1.8805×10^{-3}
2.250	1.8512×10^{-3}	1.8604×10^{-3}

Table 15. Comparison of Uncontaminated Signal and Additive Noise for $B = 0.025$

Time Length of Available Response (seconds)	RMS Value of Uncontaminated Signal	RMS Value of Additive Noise	Average Value of Additive Noise
2.25	0.14	1.65×10^{-3}	1×10^{-4}
2.0	0.15	1.75×10^{-3}	1.05×10^{-4}
1.75	0.16	1.87×10^{-3}	1.26×10^{-4}

Table 16. Comparison of Results for $B = 0.025$

(a). Sixth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Sixth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	1.65×10^{-3}	1.8×10^{-3}	11*	5.2*
2.0	1.75×10^{-3}	3.1×10^{-3}	49*	28*
1.75	1.87×10^{-3}	1.9×10^{-3}	28	12

(b). Fifth Order Attempt

Time Length of Response (seconds)	RMS Value of Additive Noise	Fifth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	1.65×10^{-3}	1.7×10^{-3}	54	37
2.0	1.75×10^{-3}	1.8×10^{-3}	48	37
1.75	1.87×10^{-3}	2.1×10^{-3}	38*	22*

(c). Fourth Order Attempt

Time Length of Response (seconds)	Value of Additive Noise	Fourth Order Attempt		
		ERMS	Per Cent Deviation	
			Max.	Ave.
2.25	1.65×10^{-3}	1.7×10^{-3}	66	43
2.0	1.75×10^{-3}	1.8×10^{-3}	13*	6.7*
1.75	1.87×10^{-3}	2×10^{-3}	28*	13*

Conclusions

The purpose of the discussions and investigations of this and the two previous chapters has been first to discuss a computer implementation of the theoretical identification procedure, and second to extend the use of the procedure as far as possible into the "uncertain region of the identification problem." The two basic conclusions which can be formed from the total investigation are:

1. With reliable data (such as the data which yielded the results of Table 5), it is possible for the identification procedure to be applied rather arbitrarily, and the results will be a fairly accurate model. By arbitrarily applying the procedure is meant that the time length of the available response is not critical as long as the available response represents the total response of the system. By a fairly accurate model is meant that all order attempts, greater than or equal to the order of the system will reduce to the same correct order model with essentially correct parameter values.

2. With unreliable data (such as the data for $B = 0.025$ which yielded the results of Table 16), the manner in which the identification problem is solved becomes more critical. That is, for different time lengths of available response, and for different order attempts the results of the identification procedure can be different order models as well as a wide range of coefficient values for models of the same order.

The manner in which the identification problem is solved is straightforward whenever fairly accurate data is available. Figure 4 illustrates the basic manner in which the procedure is applied, and

conclusion one above emphasizes the ease in which the problem is solved. For the situation where the data is unreliable, then obtaining a solution to an identification problem requires discriminating between "signal" and "noise." Frequently the best procedure in such cases is a trial and error process of applying the identification procedure to various lengths of the total response and investigating identification for models of various orders. The results are then compared before selecting the final model. The manner in which the necessary equations are selected and solved is still straight forward even when inaccurate data is used. If the user has an understanding of the identification procedure then in general it is still possible to identify the system in question, although clearly the results cannot be as good as would be the case using reliable data.

Figure 4 represents the basic manner in which the identification procedure is to be automated, and although the basic manner in which an identification problem is solved is straightforward there are several variables, termed "parameters of the implemented procedure," which exist in connection with the total identification problem. These parameters are:

1. The time length of the response required before the identification problem can be solved.
2. The order of the models for which identification is to be attempted.
3. The number of equations to be used in the iterating process.
4. The sampling interval to be used if the response of the system is to be integrated by using a numerical method.

5. The required accuracy of the data.

The success obtained when using the identification procedure depends upon all of the above factors.

The discussions of this and of all the previous chapters as well as the examples of both this and the foregoing chapter provide the user with an insight into the problems which can exist in practice. They also provide an intuitive knowledge as to how the identification problem can be solved when conditions are such that the problem lies in the "uncertain region of the identification problem."

The conclusions formed, as a result of this study, in connection with the parameters of the implemented procedure will now be given. These conclusions have been formulated during the total development and investigation of the identification procedure, and all major conclusions have been illustrated with examples throughout both this and the previous chapter.

1. A preliminary solution to the identification problem is obtained by solving k equations in k unknowns with a direct method. The equations have the form of Equation 2.36, and are selected from the total available response on an approximately linear time basis.

2. The preliminary solution should be used as starting values for the conjugate gradient method for the purpose of iterating in order to improve the solution. The equations used in conjunction with the iterating process can be the original set of equations which were solved to obtain the preliminary solution; however, if possible more equations should be used. The number of equations does not have to be large, and three to four times as many equations as unknowns is sufficient.

It is recommended that the equations used for iterating be selected on a quantized basis so that more equations will be selected during the portions of the response where it exhibits large variations.

3. Each obtained model should be checked for surplus factors and such factors removed. If the surplus factors which are removed are not identical it is recommended that the parameters of the resulting model be used as starting values for the conjugate gradient method in order to further improve the model.

4. The time length of the available response must be representative of the total impulse response of the system. That is, the available response must be sufficiently long so as to characterize all of the smaller time constants of the system. It has been stated that the essential time length of a response is the time length required for the response to decay by 99 per cent. For the example of Figure 2 this definition of essential response length was sufficient for successfully identifying the system when using reliable data; however, for unreliable data (i.e., for $B = 0.025$) approximately 99.5 per cent of the decay was required before identification was successful. The definition of essential time length will change from problem to problem depending upon the characteristics of the system being identified. In general less than 99 per cent decay is sufficient when the response of the system is either overdamped or critically damped or whenever very accurate data is available. When the response of the system is underdamped or when the data is very inaccurate then more than 99 per cent decay of the response is generally required.

5. If the exact order of the system is unknown, then the

identification attempts will take the form of first specifying an arbitrary model of order s and attempting identification. The value of s is then increased (or decreased) by one and the identification process is repeated until several attempts for different values of s all yield a model of order \hat{s} (or a model of order lower than \hat{s}). All results are compared and if all resulting models of order \hat{s} do not have essentially the same parameter values it becomes necessary to refine the solutions by using a multivariable search routine.

6. It has been found that the results of an identification attempt tend to be the lowest order model which will still closely approximate the response of the system. This means that even if the order of the system is known and identification is attempted for a model of the correct order, there is no assurance that the solution will not contain surplus factors and reduce to a lower order model. This undesirable result has occurred in practice only for cases where only a partial percentage of the total essential response length is available or whenever the measured data is very inaccurate. However, it has also been determined that when working with noisy data that higher order attempts (than the actual order of the system) will generally reduce to a model of the correct order. The problem of incorrect identification of the system order is less severe as the data becomes more accurate and/or as data is obtained from a larger per cent of the total response.

7. It is recommended that the data used in an automated identification procedure have some redundancy. (By "redundancy" is meant more than the minimum amount of data should be available.) Such redundancy is obtained whenever the response of the system is sampled and the

required integrations performed with a numerical method because then, it is possible to form an equation at each individual time at which a sample has been taken. Redundancy is particularly important when the response is contaminated with noise.

8. It is necessary to have some estimate as to the accuracy of the data used in connection with the identification procedure especially in the practical case for which the data is obtained by measurements made upon the response of the system. Furthermore whenever the required integrations are performed by using a numerical method it is necessary to estimate the accuracy of the integration process as discussed in Chapters III and IV. Thus it is recommended that the identification problem be solved using two or more sampling intervals and/or different integration rules, and that the final solutions be compared for agreement.

For the case of reliable data the identification procedure will yield a good model (for example see the results of Table 5), while for unreliable data (such as the case for $B = 0.025$), the identification procedure could yield several different models. However, it is seen that even for the case of $B = 0.025$ (the result of Table 16) that the identification procedure yielded only two models, one of which is of the correct order with approximately correct coefficient values. By using a multivariable search procedure, starting from these models, an improved model could be obtained in a straightforward fashion. However, it is also seen by inspecting the results of Table 16 that, as a typical example, the identification procedure yielded a correct order model which has a response error of 1.9×10^{-3} , while the RMS value of

the noise on the response of the system is 1.87×10^{-3} . Therefore, even if a multivariable search routine is able to produce an exact model, the minimum value of the response error is only 1.87×10^{-3} .

As discussed in Appendix E, the developed identification procedure and multivariable search routines complement each other. In employing a multivariable search routine it is necessary to specify both a model and initial estimates of the values of the parameters of the model. In this research a rather routine method has been developed for obtaining both the order of the model and a good estimate of the values of its parameters. Thus, the trial and error process inherent in using a multivariable search routine whenever a priori knowledge about the system is not available can be reduced considerably.

CHAPTER VI

MODEL OF NEUTRON DECAY WITHIN A REACTOR

The identification procedure has been applied to a practical problem of determining a model for neutron flux decay within a nuclear reactor. This chapter discusses the application of the procedure and presents the results.

Problem Statement

The problem is to determine an equation which describes the shutdown transient of a nuclear reactor at the Georgia Institute of Technology. The reactor while running at some equilibrium power level, is suddenly shut down by a reactive step and the subsequent values of the neutron level are measured. The data is obtained by counting the number of detections of neutron flux occurring in each 0.1 second interval after shutdown. The data has been recorded for approximately 40 seconds, and the value at the end of each counting interval is taken as a sample of flux decay. This gives in effect the response of the neutron flux to an impulsive input.

A plot of the experimental data consisting of 395 samples is shown in Figure 5. It should be noted that the data represents only the first portion and not the total duration of the flux decay. The decay has been estimated as lasting several hours. It should be further noted that the data is very noisy.

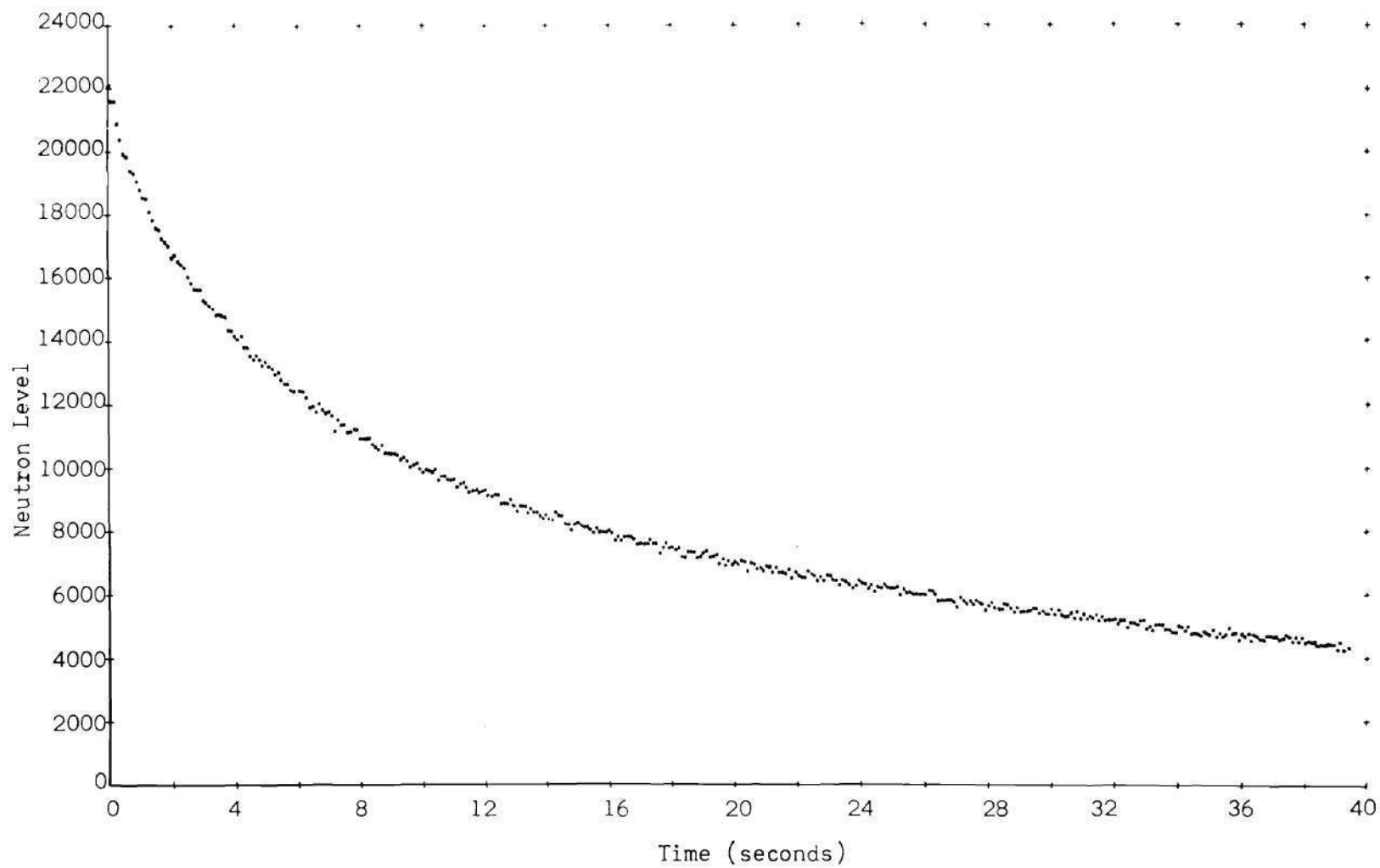


Figure 5. Plot of Neutron Flux Decay.

Solution Using the Identification Procedure

The automated identification procedure discussed in Chapters 4 and 5 can handle problems of the above type in a routine fashion. The input to the procedure includes the samples, the sampling interval, and a range of model orders for which identification is to be attempted. The procedure will then automatically perform the required integrations, form and solve the required systems of equations, and check for surplus factors. The ERMS error of each model is computed and the models as well as the ERMS errors are then printed on the line printer. A visual inspection can then be made to determine if there is agreement in the results -- such as that illustrated in Table 5 where the identification procedure was applied to reliable data.

There are two reasons why it is felt that the results of applying the identification procedure to the reactor data will not have the agreement of Table 5. These reasons are:

1. The data is very contaminated with noise, thus the results of the identification procedure will more likely be similar to the results of Table 16. There, the identification procedure was applied to the noisy data for $B = 0.025$.

2. Only a partial response is available, thus it is felt that the identification procedure will in general yield models which are of too low order.

To compensate for the fact that only a partial response is available it is necessary to apply the identification procedure in several different ways. Further, to compensate for the fact that the response is very noisy it is necessary to employ a multivariable search procedure

improve the models obtained using the identification procedure.

Method of Attack

There are a total of 395 samples, and the spacing between samples is 0.1 second. The first sample value is 22,067, and the samples decrease in value to a final sample value of 4,336. Before attempting identification, the data is normalized by dividing each sample value by 22,067. The deterministic expression of the resulting model can then be unnormalized by multiplying the expression by the normalizing expression.

The identification procedure is programmed to attempt identification for models of order three through six, and the program is to be run four times as follows:

1. The required integration performed using the Trapezoidal rule.
2. The required integration performed using Simpson's rule.
3. Every other sample value of the response removed (i.e., the sampling interval is doubled), and the required integration performed using the Trapezoidal rule.
4. Every other sample value of the response removed, and the required integration performed using Simpson's rule.

The above steps will yield a total of sixteen models, and all results are to be compared to determine if higher order attempts at identification are necessary.

Results

From a total of sixteen models determined by the identification procedure, all except three models reduced to essentially the same second order model. The best second order model in terms of having the minimum error is given by

<u>poles</u>	<u>zeros</u>	
-0.03	-0.17	(6.1)
-0.27		
Constant multiplier = 1.0		

The impulse response of the above model is

$$h(t) = 0.5438 e^{-0.03t} + 0.4622 e^{-0.27t} \quad (6.2)$$

and the error is

$$\text{ERMS} = 0.0106. \quad (6.3)$$

The four parameters of Equation 6.2 have been used as starting values for a multivariable search procedure,* and the results of using the search procedure is

$$h(t) = 0.5358 e^{-0.0264t} + 0.4318 e^{-0.2435t} \quad (6.4)$$

The response error of this model is

$$\text{ERMS} = 0.00598. \quad (6.5)$$

The three identification attempts which did not reduce to a second order model resulted in third order models. The three resulting models are

$$h(t) = 0.50953 e^{-0.02t} + 0.38863 e^{-0.18t} + 0.13948 e^{-1.36t}, \quad (6.6)$$

*The multivariable search routine used in the work of this chapter is given by Davidson (48).

$$h(t) = 0.51434 e^{-0.2t} + 0.40784 e^{-0.2t} + 0.13685 e^{-2.81t} , \quad (6.7)$$

$$h(t) = 0.45316 e^{-0.02t} + 0.36615 e^{-0.27t} + 0.24679 e^{-0.085t} . \quad (6.8)$$

The parameters of all three models have been used independently as starting values for the multivariable search procedure. The final results in all cases are identically given by

$$h(t) = 0.5004 e^{-0.02424t} + 0.4103 e^{-0.1924t} + 0.09299 e^{-1.304t} . \quad (6.9)$$

The response error for this model is

$$\text{ERMS} = 0.00435. \quad (6.10)$$

If the performance index is to be calculated using Equation 2.5, the normalizing factor is 0.419. Thus the performance index for the above model is

$$I = 0.00104. \quad (6.11)$$

The value of the performance index is particularly illustrated by this example. Since the response of both the system and the model has been normalized by a factor of 22,067 the unnormalized error of the above model is actually

$$\text{ERMS} = 95.98 . \quad (6.12)$$

Although the error is quite large the unnormalized performance index of the model is still given by Equation 6.11.

Solution Using a Multivariable Search Routine

It is known a priori that the neutron flux decay within the reactor is expressible as a sum of exponentials of the form

$$\eta(t) = \sum_{i=0}^k \alpha_i e^{-\lambda_i t} . \quad (6.13)$$

Here, α_i and λ_i represent respectively the abundance and decay constants of the i^{th} neutron group, and k represents the number of decay groups.

In general, the standard way of determining the parameters k , α_i and λ_i is to use an iterative, least-squares multivariable search procedure and a high speed digital computer. The number of terms k is not an initial constraint, but is determined by obtaining least-squares fit using different values of k . The value of k is then the number of terms which yields the optimum least-squares fit to the data.

Since only a partial response is available there is reason to believe that the identification procedure will yield a model which is of too low order. As a check, it has been decided to use the multivariable search procedure in order to determine if any better models can be found which are of higher order than the highest order model determined with the identification procedure. Employing a multivariable search procedure in this case is relatively easy since it is known a priori that the model must have the form of Equation 6.13.

The multivariable search procedure was used to find a fourth order model, and the resulting model is given by

$$h(t) = 0.5014 e^{-0.0243t} + 0.4117 e^{-0.1939t} + 0.09408 e^{-1.407} - 0.00722 e^{-600t} \quad (6.14)$$

The response error of this model is

$$ERMS = 0.00431, \quad (6.15)$$

and it is seen that the response error has been decreased in the third significant digit only.

Comparing the first three terms of Equation 6.14 with the terms of Equation 6.9 reveals that essentially the same roots are found in both equations. Further inspection of Equation 6.14 also reveals that the decay constant of the fourth term is a very large number. Since the time between samples is 0.1 second, the term involving the large time constant is negligible for $t \geq 0.1$ second. Thus this term has effectively matched the response of the model to the response of the system only at the first sample point.

The best results from the attempts at finding a fifth order model is

$$h(t) = 0.3894 e^{-0.03894t} + 0.3737 e^{-0.2205t} + 0.08394 e^{-1.69t} + 0.163 e^{-0.009992t} - 0.01002 e^{-295.18t} \quad (6.16)$$

The response error for this model is

$$\text{ERMS} = 0.00426. \quad (6.17)$$

Again, one term of this model is negligible after the first sample point (i.e., for $t \geq 0.1$) and as before, the three basic decay terms of Equation 6.9 also appear in the above fifth order model. It is also seen that a very small time constant has appeared in this model. This decay term has a value of approximately 0.01, and its appearance substantiates the concept that the actual response lasts for a period of time much greater than the time length of the available response. It is suspected that if higher order identification attempts are made that even smaller decay constants will appear.

Conclusions

The true model of the neutron decay of Figure 5 is not known; however, even though only a partial response is available the identification procedure yielded a model which has a relatively small response error -- even when compared to higher order models. This was done despite the fact that the available samples are very noisy. It is seen that the same type of results have been obtained for this problem as obtained in the examples of Chapter V -- that is, since only a partial response is available then the lower order decay roots have been approximated by a fewer number of terms.

A significant point is illustrated by the use of two identification procedures in this chapter. The computer time required to make one run (as specified under Method of Attack), using the identification procedure is approximately eight minutes, and an additional time of approximately seven minutes was required to iterate upon a resulting

third order model to improve its parameter values. However, the computer time required using only the multivariable search routine, starting from an initial guess of the parameter values, was approximately 28 minutes. The excessive time required using the search routine is due mainly to a "bad" guess as to the initial parameter values; however, this is a problem which can always be encountered whenever the guess must be made without a priori knowledge. Fortunately in this case, the form of the model is known, thus there was no wasted time because of this aspect of the problem. It should be noted that the identification program developed in this research is not the most proficiently written program with respect to running time since first the results of each step are printed on the line printer, and second the author does not have the exacting ability of a professional programmer in writing the optimum program with respect to running time required.

It should be noted that the problem of a s^{th} order model attempt reducing to a lower order model is not restricted to the identification procedure of this research, but can also occur whenever using a multivariable search routine. This fact is illustrated by an attempt at finding a third order model using only the search routine and initial guess as to the parameters. The search routine produced a second order model in this case. The same undesirable results were also obtained during an attempt at finding a fifth order model (this is an attempt different from the one which produced the fifth order model of Equation (6.16)). In this attempt, the search routine produced a fourth order model. Both of these results emphasize the previously made point that system identification can be a multimodal problem -- especially whenever working with noisy data.

CHAPTER VII

IDENTIFICATION FOR A GENERAL EXCITATION

This chapter is included to discuss the identification procedure for the situation where the system excitation is a general signal. In particular, consideration is given to the required frequency spectrum of the excitation in order to obtain successful identification.

Introduction

The general derivation of the identification procedure indicates that a system can be identified when it is excited by an arbitrary signal and both the excitation and the response can be measured. However, physical reasoning for the practical situation leads to the conclusion that there must also be requirements upon the frequency spectrum of the excitation signal.

An unknown system excited by a signal with a specified frequency spectrum is identified with respect to the particular excitation when a model is found which produces the same magnitude and phase shift in its response to this excitation. Although the identification procedure is applicable for any arbitrary excitation, because of the practical situation of having to work with noisy data there are no guarantees as to the accuracy of the derived model when the frequency spectrum of the excitation is changed. It is thus necessary to fully excite an unknown system when attempting to fit a mathematical model to the system.

A computer program has been written to investigate the above point as well as several other aspects of the identification procedure for the case of excitation by a collection of sinusoids. These aspects are

1. the effect of varying the frequency spectrum of the excitation,
2. the effect of beginning the identification process when the system has unknown but arbitrary initial conditions,
3. the effect of using different sampling intervals, and
4. the accuracy of the identification procedure in that it yields the correct order system.

Although the computer program is completely general and can be used for an extensive investigation of the procedure, the conducted investigation has been limited in scope. That is the investigation has been conducted only to formulate general conclusions about each of the above aspects of the procedure for the case of a general excitation.

Automating the Identification Procedure

As in the case of the impulse response, the automation of the identification procedure for the purpose of investigating the procedure for a general excitation involves two programs. The first program is a simulation program which generates tables of data for a specified system, and the second program is the automated identification procedure which uses the generated tables of data to form and then solve the required systems of equations.

Simulation Program

The simulation program is written for the purpose of generating the response of a specified system for a specified excitation. The excitation chosen is a sum of sine functions of different frequencies and

of constant amplitude. This allows the excitation to have a specified discrete frequency spectrum. A convenient notation for the discrete frequency components in the excitation is the following

$$w = 10^x, \text{ for } x = (x_1, x_2, x_3), \quad (7.1).$$

which connotes that x varies from x_1 in steps x_2 up to x_3 . There is a discrete frequency in the frequency spectrum of the excitation for each value of x .

The simulation program makes it possible to specify a test system along with the frequency spectrum of the excitation signal. A sampling interval and the number of required samples are also specified. The values of the excitation, of the response, and of the integrated signals are generated at the specified times. It is also possible to specify the time $t_0 \geq 0$ at which the identification procedure is applied -- that is, the system is excited at time $t = 0$; however, the start of the measuring process on the required signals does not start until time $t_0 \geq 0$. This allows the system to have arbitrary initial conditions depending upon the time t_0 at which the identification procedure is applied. All of the above variables can be specified independently.

Identification Program

The generated tables of data as well as the orders of the models for which identification is to be attempted serves as input to the identification program. The method in which the identification problem is solved is as discussed in Chapter IV. If identification is to be attempted for a model of order s , then $k \neq 3 \times s$ equations are selected in as many unknowns and solved by the method of matrix inversion. All available equations are then used in conjunction with

the conjugate gradient method to iterate upon the preliminary solution. The preliminary solution is obtained from equations which are selected on a linear time basis; however, in no case are more than $2 \times k$ equations used in the iterating process.

Outline of the Investigation

The investigation has been conducted primarily using two test systems. They are a third and a fourth order system, and a Bode plot of these systems are shown respectively in Figures 6 and 7. The third order system has the characteristics given in Table 17, while the fourth order system is the test system of Table 1. These two systems have been chosen for tests because they possess similar frequency characteristics. That is both systems have a resonance peak, and they each have approximately the same essential bandwidth*. It should be noted, however, that the third order system is not a lower order approximation of the fourth order system, and also that the resonance peak of the fourth order system is more predominant than that of the third order system.

Table 17. Characteristics of a Third Order System

Differential Equation:					
(3)	(2)	(1)	(2)	(1)	
x	$+ 7x$	$+ 20x$	$+ 24x =$	y	$+ 3y + 3.25y$
<u>poles</u>			<u>zeros</u>		
$-2 \pm j2$			$-1.5 \pm j1$		
-3					

* By essential bandwidth is meant the frequency range on a Bode plot where the system is characterized by a nonasymptotic behavior.

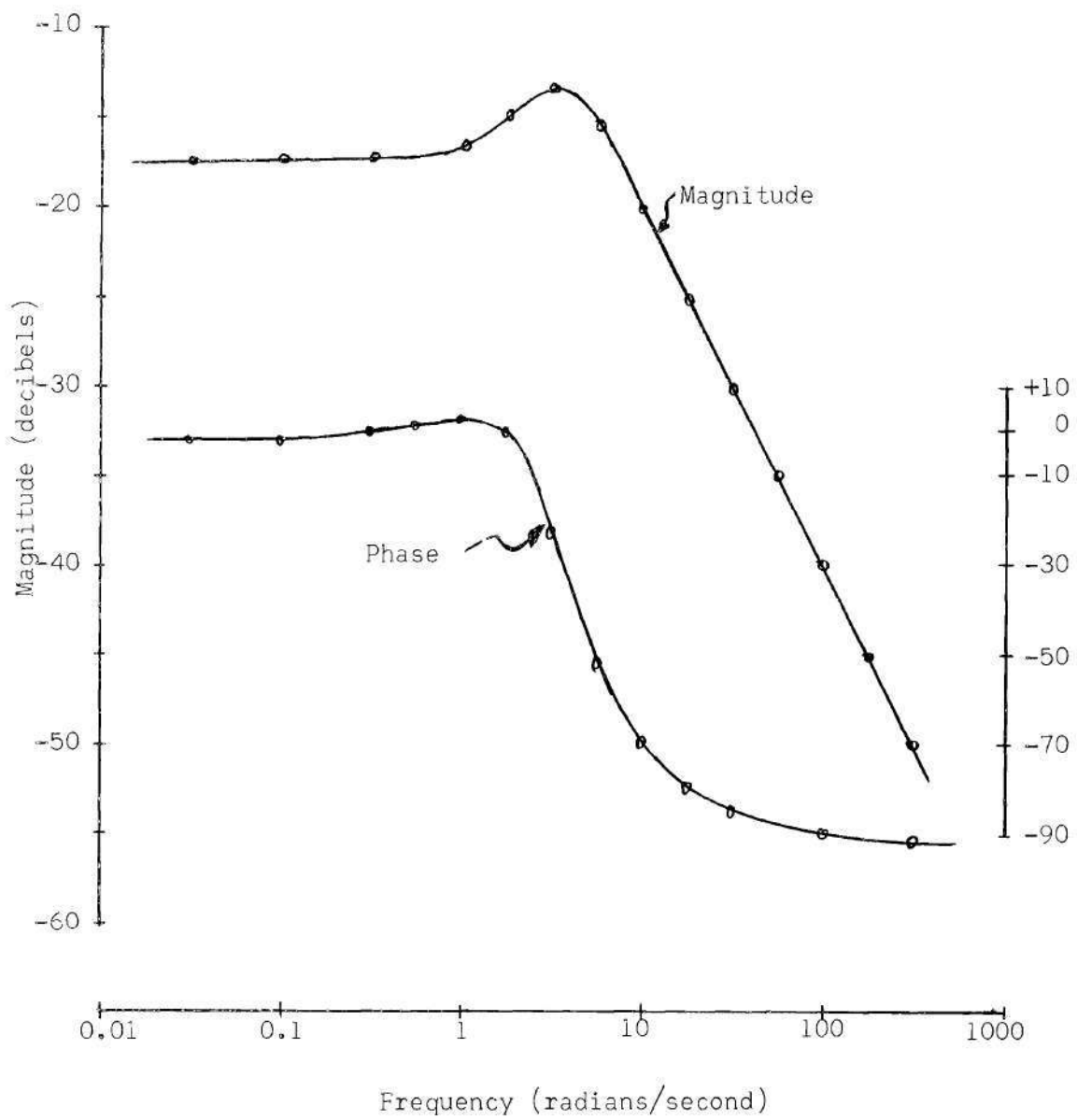


Figure 6. Frequency Characteristics of a Third Order System.

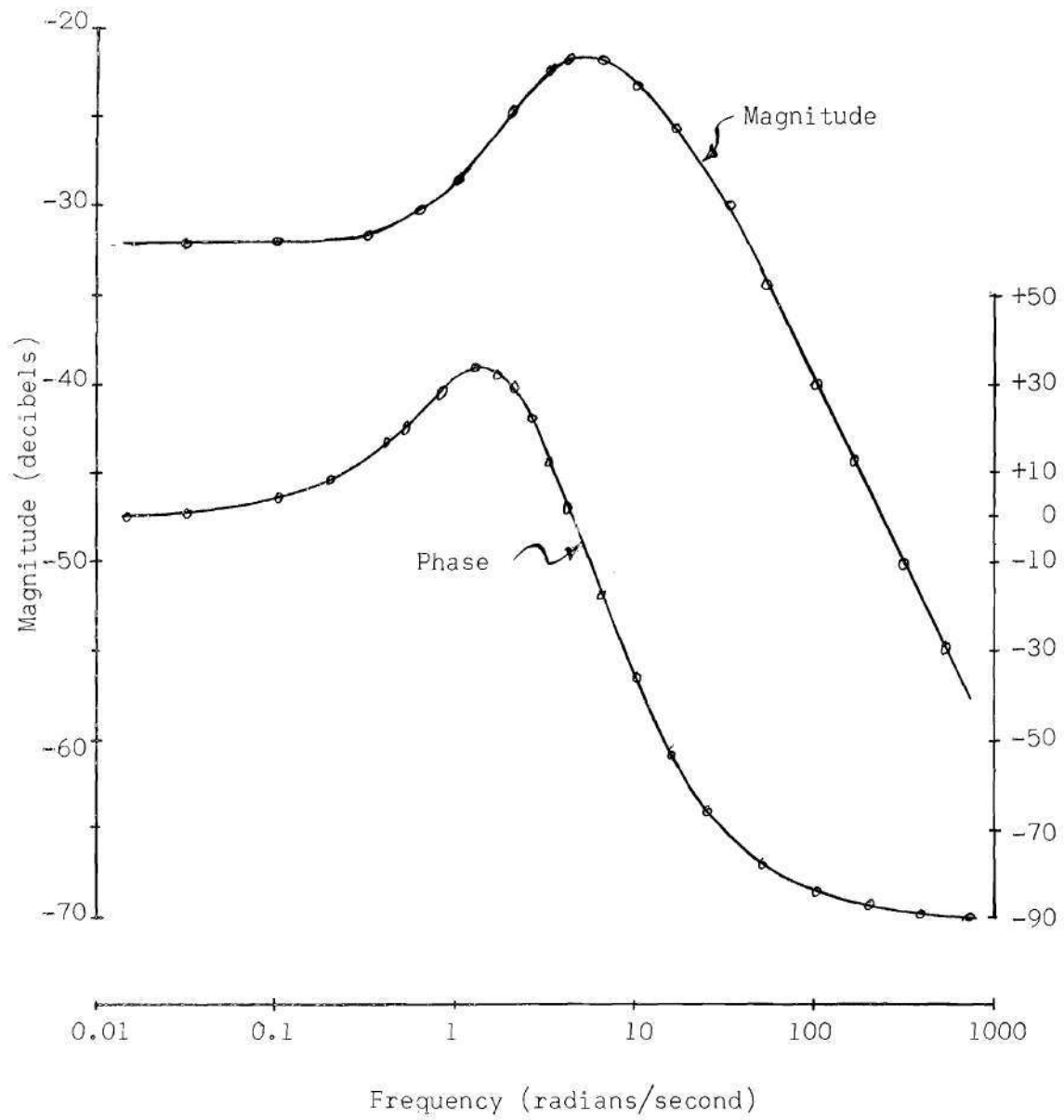


Figure 7. Frequency Characteristics of a Fourth Order System.

Identification attempts have been made when each system is excited by a variety of excitations, and Table 18 lists the various frequency spectrums of the excitations. Identification has been attempted for each excitation by using various sampling intervals ranging from one second per sample to 15 seconds per sample. In all cases, the systems were excited for a total of 20 seconds before making the required measurements.

Table 18 illustrates that the excitations consists of three different ranges of frequencies with a different number of discrete frequencies within each range. The first three excitations have a discrete frequency spectrum totally within the essential bandwidth of the test systems, while the remaining excitations also have discrete frequencies beyond the essential bandwidths. Thus, excitations four and seven contain the same discrete frequencies as excitation one as well as other discrete frequencies beyond the essential bandwidth. The same is true for excitations five and eight compared to excitation two, and for excitations six and nine compared to excitation three.

Table 18. Frequency Spectrum of Test Excitations

Excitation Number	Range of x	Number of Discrete Frequencies per Decade
1	$(-1, 0.1, 2)$	10
2	$(-1, 0.25, 2)$	5
3	$(-1, 0.5, 2)$	2
4	$(-2, 0.1, 4)$	10
5	$(-2, 0.25, 4)$	5
6	$(-2, 0.5, 4)$	2
7	$(-3, 0.1, 6)$	10
8	$(-3, 0.25, 6)$	5
9	$(-3, 0.5, 6)$	2

Results of the Identification Attempts

In all cases, the attempts at identifying the third order system for excitations one, two and three yielded a correct order model with approximately correct coefficient values. For example, the worst model in terms of coefficient values has an average absolute deviation* of only 0.73 per cent in its parameter values. On the other hand, the attempts at identifying the fourth order system for the first three excitations resulted in models which are not as accurate representations of the fourth order system. All identification results for the first two excitations resulted in a fourth order model; however, in several attempts using excitation three the results are a third order model. The resulting third order model in each case approximates the third order model given by Equation 5.10**.

In general, all fourth order models resulting from identification attempts using the first three excitations have an average deviation of less than five per cent compared to the coefficients of the true system. It should be noted, however, that in all of the above attempts, all determined models for the fourth order system have a frequency characteristic which varies from the frequency characteristic of the true system by at most 0.082 decibels and 0.32 degrees, at any of the discrete frequencies of the particular excitation which was used. This comparison also includes the third order models resulting from use of excitation three.

* Here, the deviation is computed as discussed in Chapter V, page 88.

** The third order model of Equation 5.10 resulted from an attempt at finding a lower order model approximation to the fourth order system when given the impulse response of the system.

Identification has also been attempted with each system using excitation four through nine, and in general the results of the identification attempts deteriorated as:

1. the frequency spectrum of the excitation extended beyond the essential bandwidth of the system being identified, and

2. the density of the frequency spectrum, for a specified frequency range in the excitation, decreased.

Also, in general, the results of identification attempts with the third order system are more successful than the attempts with the fourth order system, for the same excitation.

By deteriorated as used above is meant that the resulting model either reduced to a lower order approximation of the system, or if the model is of correct order then there exists a large deviation with respect to its coefficient values and the coefficients of the system it is to represent. However, it should be noted that all resulting models are close to the system being identified when compared with respect to their frequency characteristics. That is, over 90 per cent of all models have a deviation of less than 0.5 decibels and 0.5 degrees with respect to the system they are to represent.

The sampling interval was found to be a critical factor, especially as the density of the frequency spectrum of the excitation decreased or as the spectrum extended beyond the essential bandwidth of the system. For example, when the fourth order system was excited by excitation three, some sampling intervals would result in a fourth order model while some sampling intervals resulted in a third order approximate model. However, when using either excitation two or three

it was found that the length of the sampling interval was unimportant so long as the time length of the response was such as to allow the lower frequency components to complete one or two cycles.

It has also been determined that the value of any existing initial conditions within the system at the start of the identification process is unimportant. Thus, the identification procedure can be applied at any arbitrary time t_0 after the system has been excited. This result agrees with the theoretical development of the identification procedure. It should be noted that the start of the measuring of the signals is always taken as $t = 0$, and knowledge as to the value of t_0 is not required.

Conclusions and Recommendations

The principal conclusion resulting from the investigation is that the optimin excitation, for use in connection with the identification procedure, contains a frequency spectrum totally within the essential bandwidth of the system being identified. Further, the density of the discrete frequency spectrum becomes increasingly critical as the complexity of the system being identified, increases -- that is, as either the order of the system increases, or as the size of a resonant peak (or similar characteristic) increases, then the density of the frequency spectrum must increase to preserve the same accuracy.

All of the results of the investigation substantiate the previous remarks that the excitation must completely excite the system if the identification results are to be successful. An impulse and a white noise signal are two inputs which completely excite a system, but intuitively it is felt that a random noise excitation will present too

much unimportant information for the identification procedure. That is, white noise will excite the system at frequencies far beyond its essential bandwidth and thus the unimportant responses will tend to obscure or disguise the responses of the more important frequencies.

It is felt that the developed identification procedure is not as powerful for the case where the response of the system is due to a general excitation -- as compared to the case of a response due only to initial conditions within the system, or equivalently an impulse response. This is because a priori knowledge is now required as to the essential bandwidth of the system being identified. However, this restriction does not limit the generality of the identification procedure in any other way. That is, the procedure can still yield both the form of the required model as well as the value of its parameters.

Even with the above restriction, there can still be applications where the procedure is quite useful. In particular, the procedure is extremely useful when all that is desired is an approximate model which possesses essentially the same frequency characteristics as a given system. Here it may even be desirable that the model be of as low an order as possible. In this situation, the frequency spectrum of the excitation is not too critical and excellent results can be obtained.

The investigation of the identification procedure for general excitation has been restricted in scope; however, the work has proven that, first, the procedure can be implemented, and secondly, that it can be used to solve a practical problem. There are, however, several aspects of the procedure which must be investigated in further detail, and these studies are recommended in any future work with the procedure.

These recommendations are:

1. An investigation of the procedure when the excitation has a band-limited but continuous frequency spectrum.
2. An investigation of sampling on a quantized signal basis rather than a linear time basis.
3. An investigation of the effects of noise on the measured response and excitation.
4. An investigation of the use of a multivariable search routine to improve a model obtained through the application of the identification procedure.

CHAPTER VIII

CONCLUSIONS

The preceding chapters have been devoted to the development of a computational technique for the identification of a class of linear systems. The procedure is applicable for any stable, linear, time-invariant system which possesses either a low-pass or band-pass frequency characteristic. The identification procedure will determine a differential equation for the unknown system when given measured values of the excitation and response of the system. It has also been shown that a priori knowledge as to the order of the system or the value of its parameters is not required since the procedure can determine both the order of the required differential equation as well as the value of its coefficients.

The identification procedure has been developed first as a theoretical identification procedure, and it is shown that in a theoretical sense the procedure will yield exact identification -- that is, an exact model. An error analysis of the computational methods required within the identification procedure considers the practical case of inaccurate data and inexact numerical methods. In this analysis it is shown that the error in the results of an identification attempt is bounded by an expression involving the inaccuracies in the data and the errors introduced by the required numerical methods.

The identification procedure has been developed as a problem in establishing and solving a system of linear algebraic equations of the form of Equation 2.17. In the development it is shown that the only data required by the identification procedure is sampled values of the excitation and response of the system as well as integrated values of these signals. The required integration is an n^{th} order process where n is equal to or greater than the order of the system being identified. The required integrations can be performed with either analogue integrators or with numerical integration techniques.

An investigation and implementation of the identification procedure has been conducted for the case where the excitation is an impulse. The two basic conclusions are:

1. The way in which a practical identification problem is solved is straightforward. In particular, Chapter V contains examples of identification attempts as well as a conclusion section of how the procedure is to be implemented in practice.

2. The success of an identification attempt depends upon two factors: namely, the time length of the available response, and the accuracy of the available data.

For successful identification, it is necessary that the time length of the impulse response of the system be of sufficient length to characterize all of the smaller roots of the system. This will generally require that measurements be made on the response until it has decayed by approximately 99 per cent, although the exact decay value will depend on the accuracy of the data. That is, as the accuracy of the data decreases, a longer time length of the response

is required for successful identification, but as the accuracy of the data increases then a smaller time length of the response is sufficient.

By successful identification as used above is meant, first, for accurate data that the procedure will yield a correct order model with coefficients very close to their true value, and second, for very inaccurate data the procedure will still yield a correct order model; however, the coefficients will not necessarily be as true in value. For the case of very noisy data, it can become necessary to solve the identification procedure in several ways in order to obtain successful identification. This is because the identification problem lies within "the uncertain region of the identification problem." Chapter V presents examples and discusses this point in great detail.

As an example of accurate and inaccurate data, and the results of identification attempts using these sets of data, consider the results:

1. The impulse response of a fourth system is sampled, and the individual sample each have approximately five significant digits. The sampled response is integrated by the trapezoidal rule, and the resulting integrated data is estimated at having three significant digits. The identification problem is solved using this data and the resulting model is compared to the known true system. From a total of 18 identification attempts, using various time lengths of the response, all attempts yielded a correct order model. The worst model in terms of coefficient values had an average absolute percentage deviation of 1.7 per cent compared to individual coefficients of the

true system. These results are illustrated in Tables 5 and 6.

2. The above measured impulse response is contaminated by additive noise. The magnitude of the disturbance varies randomly between ± 2.5 per cent of the value of the signal at each sample time. The noisy response has approximately one or two significant digits. It is integrated as before and the identification problem is again solved using this data. From a total of nine identification attempts, four attempts yielded the correct order model, while the remainder of the attempts yielded a third order model; however, the third order model closely approximates the fourth order system with respect to its response, Table 16 presents the results of these identification attempts.

It should be noted that of the resulting fourth order models determined in two above, the worst model has an average absolute deviation of 43 per cent with respect to its coefficient values as compared to the true coefficients of the true system. The RMS error between the response of the determined model and the noisy response of the system is 1.7×10^{-3} ; however even with exact identification the minimum response error is the RMS value of the additive noise which is 1.65×10^{-3} . It is suggested within this research that a determined model be used in conjunction with a multivariable search routine in order to improve the model. This method of improving a model will be especially useful whenever the data is of such poor accuracy as to prevent the developed identification procedure from yielding a model with good parameter values. Appendix E discusses this aspect of the problem.

The practical consequence of not having representative data (where representative implies both sufficient time length of the response and sufficient accuracy in the data) is that the identification results can be a lower order model which closely approximates the system being identified. However, as illustrated by two above the identification procedure can yield excellent results even when working with inaccurate data. The accuracy of the parameters of the resulting models increases as the accuracy of the data increases.

It should be noted that the identification procedure can purposely be made to yield a model which is of lower order than the system being identified, but which will closely approximate the system with respect to its response. This is illustrated by an example in Chapter V.

While the principal emphasis of the preceding chapters has been for the case where the system excitation is an impulse, it has also been shown that the identification procedure is also applicable for the problem of identifying a system which is excited by a general signal. An investigation has been conducted in which it has been found that for successful identification, for a general excitation it is necessary that the frequency spectrum of the excitation be confined to the essential bandwidth of the system. While this now requires that a priori information be available as to the essential bandwidth of the system, it does not limit the generality of the identification procedure in any other way. That is, the procedure can still yield both the order of the required model as well as the value of its parameters. Chapter VII presents the results of the investigation of the identification procedure for the case of a general excitation, and further conclusions

and recommendations in connection with this problem are presented there.

APPENDIX A

MATHEMATICAL MODEL

This appendix is included to present a mathematical model in the form of a vector differential equation for the class of systems considered in this research as well as to derive some mathematical transformations necessary for establishment of the identification procedure. The model of a system in the form of a vector differential equation is useful for simulating the system on either analogue or digital computers. The relationships between the initial conditions of the scalar differential equation representation and the vector differential equation representation are given by transformations derived within this appendix.

State Variable Representation

A linear, time-invariant system can be described by a linear, constant coefficient differential equation of the form

$$L_n x(t) = N_m y(t), \quad (\text{A.1})$$

with initial conditions $x^{(1)}(0), x^{(2)}(0), \dots, x^{(n-1)}(0)$, where $x(t)$ represents the output variable and $y(t)$ represents the input variable. The operators L_n and N_m are defined by

$$L_n = \sum_{i=0}^n p_i \frac{d^i}{dt^i}, \quad \text{with } p_n = 1, \quad (\text{A.2})$$

$$N_m = \sum_{i=0}^m q_i \frac{d^i}{dt^i} \quad (\text{A.3})$$

For the class of systems considered within this research, m is strictly less than n in Equations A.2 and A.3.

Any differential equation of the form of Equation A.1 can be transformed to a linear vector differential equation of the form

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{y} \quad , \quad (\text{A.4})$$

$$\underline{x}(0) = \underline{x}_0 \quad .$$

This representation is called the state variable representation of the system and the choice of the transformation is not unique. That is, a particular transformation can be chosen so as to yield an optimum vector equation for a specified problem.

Webb⁽¹⁾ has shown that if m is strictly less than n in Equation A.1, then there exists a transformation such that the state variable representation contains $y(t)$ as a scalar with none of its derivatives required. This representation is especially useful for obtaining solutions by simulating the vector equation with either analogue or digital equipment.

Applying Webb's transformation yields

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{y} \quad , \quad (\text{A.5})$$

with initial conditions $\underline{x}(0)$, as the vector equivalent to Equation A.1, where

$$A = \begin{bmatrix} -a_{n-1} & 1 & 0 & \cdot & 0 \\ -a_{n-2} & 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_2 & 0 & \cdot & \cdot & 0 \\ -a_1 & 0 & \cdot & 0 & 1 \\ -a_0 & 0 & 0 & \cdot & 0 \end{bmatrix} \quad (n \times n),$$

$$B = \begin{bmatrix} 0 \\ b_m \\ \cdot \\ \cdot \\ \cdot \\ b_1 \\ b_0 \end{bmatrix} \quad \begin{matrix} (n-m-1 \text{ zeros}) \\ \\ \\ (n \times 1), \end{matrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \begin{matrix} \text{with } x_1 = x \\ \\ \\ (n \times 1) . \end{matrix}$$

Webb's transformation was derived for the general case where the system has time-varying coefficients. For the time-invariant system, the coefficients of Equations A.1 and A.5 are related by

$$a_i = p_i, \quad i = (0, 1, \dots, n-1), \quad (\text{A.6})$$

$$b_j = q_j, \quad j = (0, 1, \dots, m).$$

Transformation of Initial Conditions

Although the coefficients of Equation A.1 and Equation A.5 transform exactly, it is necessary to derive a transformation for initial conditions.

If the input variable of Equation A.1 is a unit impulse occurring at $t=0$, the response of the system can be represented as the solution of

the homogeneous equation

$$L_n x(t) = 0 \quad (A.7)$$

with initial conditions

$$\underline{x}_s(0) = \begin{bmatrix} x(0) \\ (1) \\ x(0) \\ \cdot \\ \cdot \\ (n-1) \\ x(0) \end{bmatrix},$$

where it is assumed that the system had zero initial conditions prior to its excitation by the unit impulse.

For the class of systems described by Equation A.1, there is a unique relationship between the coefficients of the operator N_m when the input is an unit impulse. This relationship can be described in matrix notation as

$$Q = P \underline{x}_s(0). \quad (A.8)$$

Here, Q is a $(n \times 1)$ vector uniquely determined by the coefficients of the N_m operator, and P is uniquely determined by the coefficients of the L_n operator. This equation will be derived at the end of this appendix, where it will be proven that P is an $(n \times n)$ nonsingular matrix; therefore, its inverse exists and it is possible to write

$$\underline{x}_s(0) = P^{-1}Q \quad (\text{A.9})$$

Similarly, if $y(t)$ is an unit impulse occurring at $t = 0$, Equation A.5 has the homogeneous solution

$$\dot{\underline{x}} = A\underline{x} \quad , \quad (\text{A.10})$$

with prescribed initial conditions

$$\underline{x}_v(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \cdot \\ \cdot \\ \cdot \\ x_n(0) \end{bmatrix}$$

Again, it is assumed $\underline{x}_v(0) = 0$ prior to the application of the unit impulse.

If Equations A.7 and A.10 represent the same system, it is possible to prove that the transformation matrix P also satisfied the relationship

$$\underline{x}_v(0) = P\underline{x}_s(0) \quad , \quad (\text{A.11})$$

and

$$\underline{x}_s(0) = P^{-1}\underline{x}_v(0). \quad (\text{A.12})$$

In the proof of Equation A.8, it will be shown that when the initial condition vector $\underline{x}_v(0)$ is established by the application of an

impulse then the vector Q is identically the vector B of Equation A.5.

It is then possible to write

$$B = Q = \underline{x}_v(0) \quad . \quad (A.13)$$

Derivation of $Q = P\underline{x}_s(0)$

Let the transfer function of a system be given by

$$\frac{X(S)}{Y(S)} = \frac{b_m S^m + b_{m-1} S^{m-1} + \dots + b_1 S + b_0}{S^n + a_{n-1} S^{n-1} + \dots + a_1 S + a_0} \quad (A.14)$$

where m is strictly less than n and the coefficients are known. If the system is in a quiescent state and $y(t)$ is a unit impulse applied at time equals zero, the impulse response of the system is then uniquely given by the inverse Laplace transform

$$x(t) = \mathcal{L}^{-1} \left\{ \frac{b_m S^m + \dots + b_0}{S^n + \dots + a_0} \right\} \quad (A.15)$$

The same system can be represented by a homogeneous differential equation and associated initial conditions

$$\begin{aligned} x^{(n)} + a_{n-1} x^{(n-1)} + \dots + a_1 x^{(1)} + a_0 x &= 0 \\ x(0), x^{(1)}(0), \dots, x^{(n-1)}(0), \end{aligned} \quad (A.16)$$

where the unknown initial conditons are due to the application of the unit impulse. The solution, of this homogeneous differential equation, which obeys the associated initial conditions must be exactly the same as that given by Equation A.15.

Taking the Laplace transform of Equation A.16 yields

$$[S^n + a_{n-1}S^{n-1} + \dots + a_0] X(S) \quad (A.17)$$

$$- [S^{n-1}x(0) + S^{n-2} \frac{(1)}{x}(0) + \dots + S \frac{(n-2)}{x}(0) + \frac{(n-1)}{x}(0)]$$

$$- [a_{n-1}(x(0) + S^{n-2} \frac{(1)}{x}(0) + \dots + S \frac{(n-3)}{x}(0) + \frac{(n-2)}{x}(0))] \\ \cdot$$

\cdot

\cdot

\cdot

$$- [a_2(Sx(0) + \frac{(1)}{x}(0))] \\ \cdot$$

$$- [a_1x(0)] = 0,$$

which can be rewritten as

(A.18)

$$\begin{aligned} [S^n + a_{n-1}S^{n-1} + \dots + a_0]X(S) &= S^{n-1}[x(0)] + S^{n-2}[a_{n-1}x(0) + \frac{(1)}{x}(0)] \\ &+ S^{n-3}[a_{n-2}x(0) + a_{n-1} \frac{(1)}{x}(0) + \frac{(2)}{x}(0)] \\ &+ S[a_1x(0) + a_3 \frac{(1)}{x}(0) + \dots + a_{n-1} \frac{(n-3)}{x}(0) + \frac{(n-2)}{x}(0)] \\ &+ [a_1x(0) + a_2 \frac{(1)}{x}(0) + \dots + a_{n-1} \frac{(n-2)}{x}(0) + \frac{(n-1)}{x}(0)]. \end{aligned}$$

If the inverse Laplace transform of Equation A.18 is to be exactly the same as the solution obtained from Equation A.15, it is necessary for the following equalities to hold

$$x(0) = b_{n-1} \quad (A.19)$$

$$a_{n-1} x(0) + x^{(1)}(0) = b_{n-2}$$

•

•

•

$$a_2 x(0) + a_3 x^{(1)}(0) + \dots + a_{n-1} x^{(n-3)}(0) + x^{(n-2)}(0) = b_1$$

$$a_1 x(0) + a_2 x^{(1)}(0) + \dots + a_{n-1} x^{(n-2)}(0) = b_0.$$

Thus it is possible to write in matrix notation

$$Q = P \underline{x}_s(0) \quad , \quad (A.20)$$

where

$$P = \begin{bmatrix} 1 & 0 & \dots & & 0 \\ a_n & 1 & & & \cdot \\ a_{n-1} & a_n & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ a_2 & a_3 & \dots & 1 & 0 \\ a_1 & a_2 & \dots & a_n & 1 \end{bmatrix} \quad (n,n),$$

$$Q = \begin{bmatrix} b_{n-1} \\ b_{n-2} \\ \cdot \\ \cdot \\ \cdot \\ b_1 \\ b_0 \end{bmatrix} \quad (n \times 1), \quad \underline{x}_s(0) = \begin{bmatrix} x(0) \\ (1) \\ x(0) \\ \cdot \\ \cdot \\ \cdot \\ (n-1) \\ x(0) \end{bmatrix} \quad \cdot$$

Since P is an $n \times n$ triangular matrix with every term along its diagonal equal to unity, it is nonsingular and thus has an inverse. It is then possible to write

$$\underline{x}_s(0) = P^{-1}Q. \quad (A.21)$$

$$\underline{\text{Derivation of }} \underline{x}_v(0) = P\underline{x}_s(0)$$

Let the homogeneous scalar differential equation of a system with known initial conditions be given by

$$L_n x(t) = 0, \quad (A.22)$$

$$\underline{x}_s(0) = \begin{bmatrix} x(0) \\ (1) \\ x(0) \\ \cdot \\ \cdot \\ \cdot \\ (n-1) \\ x(0) \end{bmatrix} \quad \cdot$$

The corresponding homogeneous vector differential equation is given by

$$\dot{\underline{x}} = A \underline{x} , \quad (A.23)$$

$$\underline{x}_v(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \cdot \\ \cdot \\ \cdot \\ x_n(0) \end{bmatrix} .$$

The solution of these equations will be respectively $x(t)$ and $x_1(t)$, and if the equations represent the same system then $x(t) \equiv x_1(t)$, for all $t \geq 0$.

The vector differential equation can be written as n-first order differential equations

$$\dot{x}_1(0) = -a_{n-1} x_1(0) + x_2(0) \quad (A.24)$$

$$\dot{x}_2(0) = -a_{n-2} x_1(0) + x_3(0)$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$\dot{x}_{n-1}(0) = -a_1 x_1(0) + x_n(0)$$

$$\dot{x}_n(0) = -a_0 x_1(0) ,$$

where $x_1(0) = x(0)$. The first equation can then be rewritten as

$$x_2(0) = x^{(1)}(0) + a_{n-1} x(0). \quad (\text{A.25})$$

In the same way, the second equation can be written as

$$\begin{aligned} x_3(0) &= x_2^{(1)}(0) + a_{n-2} x(0) \\ &= x^{(2)}(0) + a_{n-1} x^{(1)}(0) + a_{n-2} x(0). \end{aligned} \quad (\text{A.26})$$

Continuing in this way it is possible to establish a set of simultaneous equations

$$x_1(0) = x(0) \quad . \quad (\text{A.27})$$

$$x_2(0) = a_{n-1} x(0) + x^{(1)}(0)$$

$$x_3(0) = a_{n-2} x(0) + a_{n-1} x^{(1)}(0) + x^{(2)}(0)$$

.

.

.

$$x_n(0) = a_1 x(0) + a_2 x^{(1)}(0) + \dots + a_{n-1} x^{(n-2)}(0) + x^{(n-1)}(0).$$

These equations can then be written in matrix notation as

$$\underline{x}_v(0) = P \underline{x}_s(0), \quad (\text{A.28})$$

where P is defined in Equation A.20. The inverse transformation is then given by

$$\underline{x}_s(0) = P^{-1} \underline{x}_v(0). \quad (\text{A.29})$$

It should be noted that the transformations of Equations A.28 and A.29 are also valid in any arbitrary time T when it is desired to transform internal conditions between the homogeneous scalar equation and the homogeneous vector representation of the system.

$$\underline{x}_{VE}(T) = P_{\underline{x}_{SE}}(T) + \hat{P}_{\underline{y}}(T)$$

A derivation is now presented which can be used to transform between internal conditions of the scalar differential equation and the vector differential equation at any arbitrary time T . In notational form the equation is given by

$$\underline{x}_{VE}(T) = P_{\underline{x}_{SE}}(T) + \hat{P}_{\underline{y}}(T) \quad (A.30)$$

where

$$\underline{x}_{VE}(T) = \begin{bmatrix} x_1(T) \\ x_2(T) \\ \cdot \\ \cdot \\ \cdot \\ x_n(T) \end{bmatrix} \quad \begin{array}{l} \text{the internal conditions of} \\ \text{the vector representation} \\ \text{at time } T, \end{array}$$

$$\underline{x}_{SE}(T) = \begin{bmatrix} x(T) \\ x^{(1)}(T) \\ \cdot \\ \cdot \\ \cdot \\ x^{(n-1)}(T) \end{bmatrix} \quad \begin{array}{l} \text{the internal conditions of} \\ \text{the scalar representation} \\ \text{at time } T, \end{array}$$

$P = (n \times n)$ nonsingular transformation matrix defined in

Equation A.20,

$\hat{P} = (n \times n)$ matrix - to be defined,

$\underline{y}(T) = (n \times 1)$ vector to be defined.

Since P is nonsingular, the inverse transformation of Equation A.30 also exists.

If the vector differential equation

$$\dot{\underline{x}} = A \underline{x} + B \underline{y} \quad , \quad (A.31)$$

is evaluated at an arbitrary time T , it is possible to obtain the following n -first order differential equations

$$\dot{x}_1(T) = -a_{n-1} x_1(T) + x_2(T) + b_{n-1} y(T) \quad (A.32)$$

$$\dot{x}_2(T) = -a_{n-2} x_1(T) + x_3(T) + b_{n-2} y(T)$$

.

.

.

$$\dot{x}_{n-1} = -a_1 x_1(T) + x_n(T) + b_1 y(T)$$

$$\dot{x}_n = -a_0 x_1(T) + b_0 y(T)$$

Since the solution of the corresponding scalar differential equation, $x(t)$, is identically equally to $x_1(t)$ above, it is possible to rewrite the first vector equation as

$$x_2(T) = \overset{(1)}{x}(T) + a_{n-1} x(T) - b_{n-1} y(T) \quad , \quad (A.33)$$

and similarly the second equation can be rewritten as

$$\begin{aligned} x_3(T) &= x_2^{(1)}(T) + a_{n-2} x(T) - b_{n-2} y(T) \\ &= x^{(2)}(T) + a_{n-1} x^{(1)}(T) + a_{n-2} x(T) - b_{n-1} y^{(1)}(T) - b_{n-2} y(T). \end{aligned}$$

Continuing in this way it is possible to obtain the following system of simultaneous equations

$$x_1(T) = x(T) \tag{A.34}$$

$$x_2(T) = a_{n-1} x(T) + x^{(1)}(T) - b_{n-1} y(T)$$

$$x_3(T) = a_{n-2} x(T) + a_{n-1} x^{(1)}(T) + x^{(2)}(T) - b_{n-2} y(T) - b_{n-1} y^{(1)}(T)$$

.

.

.

$$\begin{aligned} x_n(T) &= a_1 x(T) + a_2 x^{(1)}(T) + \dots + a_{n-1} x^{(n-2)}(T) + x^{(n-1)}(T) \\ &\quad - b_1 y(T) - b_2 y^{(1)}(T) - \dots - b_{n-1} y^{(n-1)}(T). \end{aligned}$$

This system of equations can be written in matrix notation as

$$\underline{x}_{VE}(T) = P \underline{x}_{SE}(T) + \hat{P} \underline{y}(T), \tag{A.35}$$

where

$$P = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 \\ -b_{n-1} & 0 & & & & \cdot \\ -b_{n-2} & -b_{n-1} & & & & \cdot \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & 0 \\ -b_1 & -b_2 & \cdot & \cdot & \cdot & -b_{n-1} \end{bmatrix} \quad (n \times n),$$

$$\underline{y}(T) = \begin{bmatrix} y(T) \\ (1) \\ y(T) \\ \cdot \\ \cdot \\ \cdot \\ (n-1) \\ y(T) \end{bmatrix} \quad (n \times 1).$$

APPENDIX B

IDENTIFYING TIME-VARYING SYSTEMS

This appendix is included in order to indicate how the identification procedure of this research can be extended to include the problem of identifying time-varying systems.

Introduction

The developed identification procedure can be used to identify the parameters of a model of a time-varying system; however, there are two constraints which must be satisfied before the procedure is applicable. These constraints are

1. the form of the model must be known, and
2. the model must be both linear and time-invariant in the unknown parameters.

This requires that the model have the form

$$\sum_{i=0}^n p_i(t) \frac{d^i x}{dt^i} = \sum_{i=0}^{n-1} q_i(t) \frac{d^i y}{dt^i}, \quad (\text{B.1})$$

with initial conditions $x(0)$, $\overset{(1)}{x}(0)$, $\overset{(2)}{x}(0)$, ..., $\overset{(n-1)}{x}(0)$.

Here,

$$p_i(t) = \sum_{j=1}^{\lambda_i} a_{ij} u_{ij}(t), \quad (\text{B.2})$$

$$q_i(t) = \sum_{j=1}^{\beta_i} b_{ij} v_{ij}(t). \quad (\text{B.3})$$

For this model, a_{ij} and b_{ij} represents the unknown parameters, and $u_{ij}(t)$ and $v_{ij}(t)$ represents time varying quantities. The order of the model as well as the time-varying quantities must be known.

Identification of a time-varying system has been discussed by Diamessis⁽⁴⁶⁾; however, as mentioned previously his method of identification relays upon knowledge of the derivatives of both the excitation and response of the unknown system. The remainder of the discussion of this appendix will be devoted to deriving an equation of the form of Equation 2.17 for an example considered by Diamessis. The form of the general equation will vary with each system for which identification is attempted, thus a completely general approach is not considered practical. This example is intended to illustrate a general method which can be used to form the required system of equations in a wide variety of cases.

Example

The example is a system which can be described by the well known Mathieu equation, and the model is given by

$$\frac{d^2 x}{dt^2} + (w^2 + \epsilon \cos t)x = y, \quad (\text{B.4})$$

with initial conditions $x(0)$ and $\dot{x}^{(1)}(0)$. Here, w^2 , ϵ , and the initial conditions are unknown. It is easier to derive the required general form of the equation from the vector differential equation representation

of the model, thus it is necessary to first transform the scalar differential equation to the vector differential equation representation (as discussed in Appendix A). The vector representation of Equation B.4 is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -(w^2 + \epsilon \cos t) & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} y_1, \quad (\text{B.5})$$

with initial conditions

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}.$$

For this particular example, the coefficients transform directly between scalar and vector representation; however, in general this direct transformation will not be true. Webb⁽¹⁾ has derived the general transformation between scalar and vector differential equation representation for systems with time-varying coefficients, and the reader is referenced to his work for the required form of the transformation.

It must be remembered that the initial conditions of Equation B.1 and B.5 are not the same. In Appendix A, a transformation between initial conditions for the scalar and the vector differential equation representations was derived for the case where the coefficients are time-invariant; however, this transformation is in general no longer applicable for time-varying models. It is possible to derive the proper transformations for time-varying models, and the work of Appendix A can be extended to the time-varying situation.

In examining the vector differential equation representation of

the model as given by Equation B.5 it is seen that

$$x_2(t_j) = x_2(0) + \int_0^{t_j} y \, dt_j - w^2 \int_0^{t_j} x_1 \, dt_j \quad (\text{B.6})$$

$$- \varepsilon \int_0^{t_j} \cos(t) x_1 \, dt_j,$$

$$x_1(t_j) = x_1(0) + \int_0^{t_j} x_2 \, dt_j. \quad (\text{B.7})$$

Substituting Equation B.6 into Equation B.7, and simplifying then yields

$$x_1(t_j) = \sum_{i=0}^1 x_{i+1}(0) \frac{(t)^i}{(i)!} + w^2 A_2(-x, t_j) \quad (\text{B.8})$$

$$+ \varepsilon A_2(-x \cos(t), t_j) + B_2(y, t_j),$$

where the A and B operators are defined respectively by Equations 2.14 and 2.15. For the case where the response is due only to initial conditions within the system, Equation B.8 becomes

$$x(t_j) = \sum_{i=0}^1 x_{i+1}(0) \frac{(t)^i}{(i)!} + w^2 A_2(-x, t_j) \quad (\text{B.9})$$

$$+ \varepsilon A_2(-x \cos(t), t_j).$$

Equations B.8 and B.9 depend only upon the response and excitation and the integrated values of these signals. Thus, it is now possible to form a system of linear algebraic equations. The

procedure developed in this research can be used in obtaining a solution to these equations, and the solution will yield the parameters of the model.

APPENDIX C

NORMED ERROR ANALYSIS

This appendix presents an analysis of the effect of errors in the coefficients and right hand member of a system of equations. The final results is a bound upon the relative errors of the solution. The final results also illustrates the effect of ill-conditioning on the system of equations.

Norms of Vectors and Matrices

It is useful to have a single number which gives an overall assessment of the size of a vector or matrix and which plays the same role as the modulus in the case of a complex number. The following discussion defines the norms of vectors and matrices used in this discussion, and gives the properties of these norms. The properties are given without motivation; however, all required proofs are given by Fadeeva* and Wilkinson**.

The norm of a vector \underline{x} will be denoted by $||\underline{x}||$ and will satisfy the relation

$$||\underline{x}|| \geq 0, \text{ and } ||\underline{x}|| = 0 \text{ only if } \underline{x} = 0, \quad (\text{C.1})$$

$$||k \underline{x}|| = |k| ||\underline{x}||, \text{ where } k \text{ is in general a complex number,}$$

$$||\underline{x} + \underline{y}|| \leq ||\underline{x}|| + ||\underline{y}||.$$

* Fadeeva (21), pp. 54-60.

** Wilkinson (22), pp. 80-81.

There are three definitions of $||\underline{x}||$ which will be considered in the discussion to follow. All three definitions satisfy the relationships of Equation C.1, and are

$$||\underline{x}||_1 = \max_i |x_i|, \quad (C.2)$$

$$||\underline{x}||_2 = \sum_{i=1}^n |x_i|, \quad (C.3)$$

$$||\underline{x}||_3 = \left\{ \sum_{i=1}^n |x_i|^2 \right\}^{1/2}. \quad (C.4)$$

Here, x_i is the i^{th} element of the vector, and there are a total of n -elements. The following inequalities are satisfied by the above norms

$$||\underline{x}||_1 \leq ||\underline{x}||_2 \leq n ||\underline{x}||_1, \quad (C.5)$$

$$||\underline{x}||_1 \leq ||\underline{x}||_3 \leq \sqrt{n} ||\underline{x}||_1.$$

Similarly, the norm of a matrix A will be denoted by $||A||$ and will satisfy the relationships

$$||A|| \geq 0, \text{ and } ||A|| = 0 \text{ only if } A = 0, \quad (C.6)$$

$$||k A|| = |k| ||A||, \text{ where } k \text{ is in general a complex number,}$$

$$||A + B|| \leq ||A|| + ||B||,$$

$$||A B|| \leq ||A|| ||B||.$$

The norm of a matrix is compatible with a vector norm if for any matrix A and any vector \underline{x} , the following inequality is satisfied

$$||A \underline{x}||_i \leq ||A||_i ||\underline{x}||_i \quad (C.7)$$

where i defines the particular vector norm considered. This requires that the matrix norm be given by

$$||A||_i = \max \frac{||A \underline{x}||_i}{||\underline{x}||_i}, \quad (C.8)$$

for all possible matrices and vectors of the same size.

Subject to the requirement of Equation C.8, the matrix norms corresponding to the vector norms of Equations C.2, C.3, and C.4 are respectively,

$$||A||_1 = \max_i \left\{ \sum_{j=1}^n |a_{ij}| \right\}, \quad (C.9)$$

$$||A||_2 = \max_j \left\{ \sum_{i=1}^n |a_{ij}| \right\}, \quad (C.10)$$

$$||A||_3 = \left\{ \text{maximum eigenvalue of } A'A \right\}^{1/2}.$$

where A' denotes the transpose of A .

There is another matrix norm which is consistent with the vector norm $||\underline{x}||_3$. This is the Euclidean norm defined by

$$||A||_E = \left\{ \sum_i \sum_j |a_{ij}|^2 \right\}^{1/2}. \quad (C.12)$$

In strictly mathematical work, the matrix norm $||A||_3$ is used in conjunction with $||x||_3$; however, the Euclidean norm has the advantage over the norm $||A||_3$ in that it is easier to compute.

The following inequalities exist among the matrix norms

$$||A||_3 \leq ||A||_E \leq n^{1/2} ||A||_3, \quad (C.13)$$

$$|| |A| ||_3 \leq || |A| ||_E = ||A||_E \leq n^{1/2} ||A||_3, \quad (C.14)$$

$$||A||_3^2 \leq ||A'A||_1 \leq ||A'||_1 ||A||_1 = ||A||_2 ||A||_1, \quad (C.15)$$

where $|A|$ denotes the matrix composed of the absolute value of the elements of A .

Ill-Conditioned Matrices

When making estimates of errors in matrix processes, it is found that the chief factor limiting the accuracy that can be obtained is the extent of the "ill-conditioning" of the matrix involved. It is characteristic of ill-conditioned sets of equations that small percentage errors in the coefficients can lead to large percentage errors in the solution. It is often stated that ill-conditioned matrices are ones which have small determinant values. While ill-conditioning and small determinant values tend to go together, the value of the determinant is not in general a representative conditioning indicator. A better conditioning indicator in terms of norms is derived in this section.

A system of n -algebraic equations are to be solved, and for notational purpose the system of equations will be given by

$$(A + E) (\underline{c} + \underline{h}) = (\underline{b} + \underline{k}) , \quad (C.16)$$

where the true system of equations is given by

$$A \underline{c} = \underline{b} \quad (C.17)$$

Here, E is an error matrix added to the true coefficient matrix A , and \underline{k} is an error vector added to the true right hand side vector b . The true solution of Equation C.17 is given by \underline{c}' ; however, the solution of Equation C.16 has additive errors denoted by \underline{h} . All norms used in the following discussion will be of type three unless otherwise specified.

If the determinant of A denoted as $\det(A)$ is zero then Equation C.17 fails to have an unique solution, and small values of $\det(A)$ might in fact indicate ill-conditioning. Even if A is not singular (as will be assumed), $A + E$ can be singular if E is not restricted. Writing

$$(A + E) = A(I + A^{-1}E), \quad (C.18)$$

then $A + E$ is non-singular provided $I + A^{-1}E$ is nonsingular. Here, A^{-1} denotes the true inverse of A , and I is the identity matrix.

A necessary and sufficient condition for $(I + A^{-1}E)^{-1}$ to exist is that

$$\|A^{-1}E\| < 1, \quad (C.19)$$

for any norm except the Euclidean*. Assuming that this condition exists,

*Fadeeva (21), pp. 60-62.

it is now possible to write

$$A^{-1} E = F. \quad (C.20)$$

Equation C.16 can be rewritten as

$$(A + E) \underline{h} = \underline{k} - E \underline{c} , \quad (C.21)$$

and the value of \underline{h} is now given by

$$\begin{aligned} \underline{h} &= (A + E)^{-1} \underline{k} - (A + E)^{-1} E \underline{c} , \\ &= (I + F)^{-1} A^{-1} (\underline{k} - E \underline{c}). \end{aligned} \quad (C.22)$$

writing

$$G = (I + F)^{-1} \quad (C.23)$$

then

$$I = G + FG , \quad (C.24)$$

and based upon norms it is possible to write

$$1 \geq ||G|| - ||F|| ||G|| . \quad (C.25)$$

Hence since $||F|| < 1$, and

$$||G|| \leq \frac{1}{1 - ||F||} , \quad (C.26)$$

writing

$$\underline{h} = GA^{-1} (\underline{k} - E \underline{c}) , \quad (C.27)$$

and taking the norm of both sides of this equation yields

$$||\underline{h}|| \leq ||G|| ||A^{-1}|| ||\underline{k}|| + ||G|| ||A^{-1}|| ||E|| ||\underline{c}||. \quad (C.28)$$

From Equation C.17 it is also possible to write

$$||\underline{b}|| \leq ||A|| ||\underline{c}||, \quad (C.29)$$

then

$$||\underline{c}|| \geq \frac{||\underline{b}||}{||A||}. \quad (C.30)$$

Using Equation C.30 it is now possible to divide Equation C.28 by $||\underline{c}||$ (or by some number smaller than $||\underline{c}||$) to obtain

$$\frac{||\underline{h}||}{||\underline{c}||} \leq \frac{||G|| ||A^{-1}|| ||\underline{k}|| ||A||}{||\underline{b}||} + ||G|| ||A^{-1}|| ||E|| \quad (C.31)$$

Substituting Equation C.26 for $||G||$ in Equation C.31 yields

$$\frac{||\underline{h}||}{||\underline{c}||} \leq \frac{||A|| ||A^{-1}|| \frac{||\underline{k}||}{||\underline{b}||}}{1 - ||A^{-1}|| ||E||} + \frac{||A^{-1}|| ||E||}{1 - ||A^{-1}|| ||E||} \quad (C.32)$$

Multiplying and dividing the expression $||A^{-1}|| ||E||$ by $||A||$ it is possible to obtain the more uniform expression

$$\frac{\| \underline{h} \|}{\| \underline{c} \|} \leq \frac{\|A\| \|A^{-1}\| \frac{\|k\|}{\|b\|}}{1 - \|A\| \|A^{-1}\| \frac{\|E\|}{\|A\|}} + \frac{\|A\| \|A^{-1}\| \frac{\|E\|}{\|A\|}}{1 - \|A\| \|A^{-1}\| \frac{\|E\|}{\|A\|}}. \quad (C.33)$$

Equation C.33 is a general inequality which bounds the relative error in the solution by the relative errors in the system of equations. It is impossible to compute this bound in practice since the error matrices and vector and the true system matrices and vectors are not known separately. However, the inequality does show an aspect of ill conditioning. If the number defined by

$$\eta = \|A\| \|A^{-1}\|, \quad (C.34)$$

is large, then the relative errors in the system of equations are reflected by a large multiplier. Wilkinson⁽²²⁾ gives the name "spectral conditioning number" to the quantity defined by Equation C.34. The most general norm used to compute the spectral conditioning number will be $\|\cdot\|_3$.

APPENDIX D

SOLVING A SYSTEM OF LINEAR ALGEBRAIC EQUATIONS

This appendix is included to present typical results of tests of different numerical methods for solving a system of linear algebraic equations, and is a supplement to the discussion of Chapter IV. The data required to form the systems of equations is obtained from a table of data prepared especially for the process of implementing the identification procedure as explained in Chapter IV.

All data given within this appendix has been obtained as print-out from a computer program. All results were obtained to five digits excluding leading zeros; however, for convenience trailing zeros have been omitted in presenting the results. Also for convenience, the values given for the norms of the residue vectors have been rounded to two or three digits since only their relative size is important.

Solutions Obtained by Direct Methods

The table of test data for the fourth order test system, discussed in Chapter IV, has been used to establish a system of equations of size (8×8) . The equations are selected on a linear time basis from the complete set of equations, and the data is such that no exact solution exists for the complete set of equations -- that is, no solution exists which will result in a residue vector with all zero elements. If the response of the fourth order test system could be sampled and processed in an ideal manner the resulting system of

equations would have a solution namely

$$[1, 7, 16, 10, 19, 118, 320, 400]. \quad (\text{D.1})$$

For comparative purposes each of the direct methods have been used to solve the established system of equations. The resulting solutions are as follows.

1. Matrix inversion:

$$[1.0, 7.0089, 16.027, 10.024, 19.009, 118.11, 320.38, 400.56]. \quad (\text{D.2})$$

2. Gause elimination:

$$[1.0, 7.0111, 16.032, 10.028, 19.001, 118.14, 320.49, 400.72]. \quad (\text{D.3})$$

3. Crout reduction:

$$[1.0, 7.0111, 16.032, 10.028, 19.011, 118.14, 320.49, 400.72]. \quad (\text{D.4})$$

The sum of the square of the elements of the residue vector for the above solutions are given respectively by

$$||x||_{\text{MI}} = 4.62 \times 10^{-15}, \quad (\text{D.5})$$

$$||x||_{\text{GE}} = 1.22 \times 10^{-15},$$

$$||x||_{\text{CR}} = 7.74 \times 10^{-15}.$$

Comparison of each of the above obtained solutions with the true solution as well as comparing the three solutions with each other illustrates that either of the direct methods yields satisfactory results. It is of interest to note that the norm of the residue vector using the established system of equations and the true solution of

D.5 is 1.9×10^{-3} . This answer is very large compared to the norms of D.5, and also illustrates that the established system of equations are not exact.

Solution Obtained by Iterative Methods

The same established system of equations is now used to test both the conjugate gradient method and the method of steepest descent. In each case, the starting values are assigned as zero, and a total of 16 iterations are performed. In the case of the conjugate gradient method, the iterative process is stopped after eight iterations, and the resulting solution used as starting values for the remaining iterations.

After eight iterations, the conjugate gradient method yielded

$$[0.99845, -0.72533, 0.8553, -0.49186, 11.267, 9.81, -5.408, 1.5578], \quad (\text{D.6})$$

and this solution has a residue vector which has a norm of

$$||r|| = 8 \times 10^{-6}. \quad (\text{D.7})$$

Eight more iterations on the values of D.6 reduced the norm of the residue vector to

$$||r|| = 1.52 \times 10^{-6}, \quad (\text{D.8})$$

and resulted in only slight changes in the values of the solution. The model resulting from the solution given by D.6 is unstable because it has two poles and three zeros in the right hand plane of the complex frequency plane.

Not all the tests conducted using the method of conjugate gradient produced such inaccurate solutions; however, in general,

results were unsatisfactory. The above example readily illustrates the discussion of Chapter III on ill-conditioned matrices. The norm of the residue vector of D.8 is relatively small when compared to D.4; however, the obtained solution is far from being a true least squares solution.

The method of steepest descent has been found to be very unsatisfactory in all tests. Using the above example, the method yielded a solution after eight iterations which has a relatively large residue vector. The norm of this vector is

$$||r|| = 2.5, \quad (D.9)$$

and eight more iterations only reduced the norm to

$$||r|| = 2.14, \quad (D.10)$$

with only slight changes to the preliminary solution. The final solution is given by

$$[0.24497, -0.23873, 0.033121, 0.10036, 0.013766, 0.015663, \quad (D.11) \\ 4.6681 \times 10^{-4}, -4.173 \times 10^{-3}],$$

and it bears no resemblance to the solution obtained by the direct methods. This solution also yields an unstable model, and is typical of the results obtained whenever using the method of steepest descent.

Solution of the Normal Equations

The normal equations are formed from the previous example, and these equations are solved by matrix inversion. The resulting solution is given by

$$[1.0, 2.5194, 3.7676, 1.4448, 14.605, 51.078, 91.639, 62.312] \quad (D.12)$$

and the norm of the residue vector of this solution is

$$||r|| = 3.86 \times 10^{-5}. \quad (D.13)$$

These values were then used as starting values for the conjugate gradient method in conjunction with the regular system of equations. After eight iterations, the solution is given by

$$[1.0, 2.5194, 3.7673, 1.448, 14.605, 51.078, 91.639, 62.312] \quad (D.14)$$

and the norm of the residue vector

$$||r|| = 1.06 \times 10^{-13}. \quad (D.15)$$

Again, these solutions bear no resemblance to the solutions given by the direct methods; however, its residue vector has a norm very close to the norms of D.4. Again, this is an example of the effects of ill-conditioning.

The model given by D.14 is a stable fourth order model which has an impulse response very close to the response of the true fourth order system. However, in general, the solution obtained by solving the normal equations were unstable. In no tests did the normal equations yield a solution which was as good as a solution which was obtained by solving the regular equations by a direct method. When the number of equations are increased, the results of solving the normal equations are even poorer than when the same number of equations as unknowns are solved.

Iterating With T Equations in k Unknowns

Both the conjugate gradient method and the method of steepest descent can be used to solve a system of T equations in k unknowns, for $T \geq k$. In theory, it is then possible to use either of these methods in place of forming and solving the normal equations. Both of the iterative methods have yielded poor results in all tests where the starting values are assigned as zero; however, it has been found that if the starting values are at best poor approximations to the true solution, the conjugate gradient method will rapidly converge to a true least squares solution. As an example, a system of (47 x 8) equations are established for the same example, and starting values are assigned as

$$[0, 0, 0, 0, 19.011, 118.14, 320.49, 400.72] \quad (D.16)$$

The norm of the residue vector for this starting value is

$$||r|| = 7 \times 10^4. \quad (D.17)$$

After eight iterations, the solution is given by

$$[0.99974, 7.0211, 15.983, 10.03, 19.006, 118.28, 319.84, 400.2] \quad (D.18)$$

and the norm of the residue vector is given by

$$||r|| = 1.74 \times 10^{-6}. \quad (D.19)$$

Eight more iterations yielded only slight changes in the solution, and the final solution is given by

$$[1, 6.9989, 16.01, 10.013, 18.996, 117.97, 319.99, 400.16] \quad (D.20)$$

The norm of the residue vector is given by

$$||r|| = 4.3 \times 10^{-10}. \quad (D.21)$$

Comparison of these solutions with the solutions obtained by the direct methods illustrates the accuracy and the rapid convergence of the conjugate gradient method when given good starting values. It should be noted that the norm of the residue vector of D.21 is relatively small compared to the norm given by D.5 although this norm involves 47 element values while the norms of D.5 involves only eight element values.

Although the data given in this appendix are to five digits, the B5500 computer has a word length of approximately eleven digits. This is the reason that D.16 and D.18 are essentially the same although there is a large difference in the norms of their residue vectors -- that is, the last eight iterations effected mainly the last digits of the preliminary solution.

If a preliminary solution is first obtained by solving a system of equations of size $(k \times k)$, this preliminary solution can then be used as a starting value for the conjugate gradient method in conjunction with a larger system of equations. As an example, the same system of 47 equations are given the starting values obtained by matrix inversion (see D.2). After eight iterations, the answer is given by

$$[1.0, 7.0091, 16.027, 10.024, 19.006, 118.11, 320.38, 400.56] \quad (D.22)$$

and the norm of the residue vector is

$$||r|| = 4.2 \times 10^{-11}. \quad (D.23)$$

The fact that the preliminary solution is changed only slightly indicates that it is a good approximate solution for the total set of equations. Again, the norm of D.23 is computed for 47 element values, while the norms of D.5 are for eight elements. The results also indicate that iterating improves the total accuracy of the solution obtained by matrix inversion. This is in effect reducing the roundoff error which is inherent in the direct methods.

APPENDIX E

IDENTIFICATION USING MULTIVARIABLE SEARCH

This appendix is included to briefly discuss multivariable search procedures and their possible use in conjunction with the identification procedure of this research. The possible application of multivariable search techniques, as discussed in this appendix, is for the case where a system is being identified based upon its measured impulse responses; however, the theory is equally applicable for the case of a more general excitation.

Multivariable Search

The identification procedure presented in Chapter II is based on establishing and solving a system of linear equations for the least squares solution. An alternate method for obtaining a solution to Equation 2.17 is through the use of multivariable search. In multivariable search, rather than solve a system of equations for the least squares solution, it is necessary to specify a model, and to establish a functional which is to be minimized. For this identification problem, the desired functional is the ERMS error as given by Equation 2.4; however, computationally it is easier to minimize a functional of the form

$$J[a_1, a_2, \dots, a_{2S}] = \sum_{j=1}^T [x(t_j) - z(t_j)]^2, \quad (E.1)$$

where

s - order of model,

$x(t_j)$ - value of j^{th} sample of measured response,

$z(t_j)$ - value of response of model at j^{th} sample time

T - number of samples,

α_i - parameters of model.

Comparing Equations E.1 and 2.4 shows that minimizing $J[\cdot]$ will automatically minimize ERMS.

The methods for locating the extremum of a function of several variables in general rely upon iterative procedures which will converge to the true extremum. There are numerous search procedures available, and each variation has a particular characteristic of convergence depending upon some specified form of the function being minimized. Woodrow (37) compares several gradient methods both theoretically and with actual problems, and concludes that his investigation has essentially solved the problem of locating the extrema of unimodal functions*.

Multidimensionality makes a unimodality assumption, about a particular function implausible since it becomes difficult to believe in unimodal response surfaces as the number of dimensions increases. Multimodality not only increases the possibility that a search procedure will locate a local minimum rather than the true minimum, it also adds the problem that a saddle point can be accepted as a minimum. However, if a correct model is available along with good initial estimates of the values of the parameters, it is possible to have essentially a unimodal functional. For a multivariable search routine to be practical,

* For other discussions of multivariable search procedures, the reader is also referred to Powell (41) and Wilde (42), and the bibliographies of these references.

approximate values of the parameters, as well as the form of a model must be specified along with the functional which is to be minimized.

It is not difficult to believe that system identification is a multimodal process. Examples are given in both Chapter V and VII of models which have responses very close to the response of the true system; however, the models are not correct. Examples are given for both the case when the model is of different order than the true system, and the case when the model is of the same order as the system, but of different coefficient values.

A comparison of the identification procedure of this research and multivariable search procedures shows that these two procedures complement each other. Search techniques require both a model and initial estimates of the parameters of the model before there is assurance that the functional will be truly minimized. On the other hand, the identification procedure of this research will yield both a model and initial estimates of the parameters. Thus, multivariable search techniques can be employed to improve a model constructed from noisy data by the routine identification procedure of this research.

Improving a Model

In conjunction with specifying the functional of Equation E.1, it is also necessary, in general, to be able to calculate the gradient of the functional whenever a multivariable search routine is employed. The gradient is simply a vector in which the $k^{th}+1$ element is the partial derivative of the functional with respect to the $k^{th}+1$ parameter. This element is then given by

$$\frac{\partial J}{\partial a_k} = \sum_{j=1}^T [x(t_j) - z(t_j)](-2) \frac{\partial z(t_j)}{\partial a_k} . \quad (E.2)$$

It is thus necessary to specify a model so that the variation of the response of the model with respect to its parameters can be calculated.

The first possible model to be used in the iteration is defined by

$$z(t_j) = \sum_{i=0}^{s-1} a_{i+1} A_{s-1}(-x, t_j) + \sum_{i=s}^{2s-1} a_{i+1} \frac{(t_j)^i}{(i)!} , \quad (E.3)$$

where $A_{s-1}(-x, t_j)$ represents the integrated values of the measured responses. There is a maximum amount of improvement which can be gained using this model since the $A_{s-1}(-x, t_j)$, which are constructed by integrating noisy data, are subject to uncertainty. In reality, use of the above model in conjunction with a multivariable search technique is equivalent to using the conjugate gradient method in conjunction with a larger system of equations. Thus, any improvement in the model which can be obtained by employing search techniques has actually been obtained within the identification procedure of this research.

A second possible method is to use the model of Equation E.3 with $A_{s-1}(-z, t_j)$, where $z(t)$ represents the actual response of the model for the given estimate of parameters of the model. As with the first method, the preliminary estimates of the parameters are found using the identification procedure outlined in this research. This second method allows the model to be improved even when the measured response is noisy, since

the response of the model as well as the partials of the functional are in this case subject to less error than the corresponding quantities in the first case.

If the second method is used, it is necessary to generate the response of the model after each iteration and recalculate the partials of the response of the model. This requires a lot of computational time; however, if the preliminary model has small error, and if the iterative scheme converges rapidly, then the process will have to be repeated a minimum number of times. It is necessary, however, to use this second method if any significant improvement in the model is to be gained, especially if the measured response is noisy.

A third method can be used when a model is being established, from a measured impulse response. This method is simply to obtain a time expression for the impulse response of the preliminary model and use this numerical expression in place of Equation E.3. This method has the greatest simplicity, since it is not necessary to integrate the response of the model. The desired partials can be determined directly from the mathematical expression. The parameters of the functions will then be the coefficients and decay term of the deterministic expression.

There is one disadvantage to the third method that is not encountered in the second method. This disadvantage occurs when the poles of the model do not agree in type with the poles of the system. The types of poles are established once the transformation is made from coefficients and initial conditions to a deterministic expression; hence the roots of the model can only be changed in value and not in form during the multivariable search. However, when the model is left in the form

of Equation E.2, the forms of the poles are not restricted. Hence if the preliminary model is such a poor approximation that its poles are not of the same type as the poles of the system, multivariable search must be conducted using the second rather than the third method.

BIBLIOGRAPHY

1. Webb, R. P., Synthesis of Measurement Systems, Ph.D. Thesis, Georgia Institute of Technology, May, 1961.
2. Webb, R. P., and Hammond, J. L. Jr., Generation and Measurements of Nonstationary Random Processes, Electrical Engineering Experiment Station, Georgia Institute of Technology, Technical Note No. 3, Project No. A-588, July 22, 1964.
3. Plackett, R. L., Principles of Regression Analysis, Oxford University Press, London, England, 1960.
4. Luenberger, D. G., "Determining the State of a Linear System with Observers of Low Dynamic Order," Institute in Engineering-Economic Systems, Stanford University CCS-1, July, 1963.
5. Mishkin, E. and Braun, L. Jr., Adaptive Control Systems, McGraw-Hill Book Company, Inc., New York, N. Y., 1961.
6. Cooper, G. R., "Limits on the Identification Time for Linear Systems," Purdue University, School of Electrical Engineering, Report PRF 2358, Lafayette, Indiana, February 1, 1961.
7. Hsieh, H. C., "The Least Squares Estimation of Linear and Non-Linear System Weighting Function Matrices," Information and Control, vol. 7, 1964, pp. 84-114.
8. Lee, Y. W., "Application of Statistical Methods to Communication Problems," M.I.T. Research Laboratory of Electronics, Technical Report No. 181, September, 1950, pp. 25-28.
9. Lee, Y. W., Statistical Theory of Communications, John Wiley and Sons, Inc., New York, N. Y., April, 1963.
10. Anderson, G. W., Buland, R. N., and Cooper, G. R., "Use of Cross-correlation in an Adaptive Control System," Proceedings of the National Electronics Conference, vol. 15, October, 1959.
11. Levin, M. J., "Optimum Estimation of Impulse Response in the Presence of Noise," Institute of Radio Engineers Transactions on Circuit Theory, CT-7, March, 1960, pp. 50-56.
12. Kerr, R. B., and Surber, W. H., Jr., "Prediction of Impulse-Response Identification Based on Short, Normal Operation Records," I.R.E. Transactions on Automatic Control, AC-6, May, 1961, pp. 173-182.

13. Turin, G. L., "On the Estimation in the Presence of Noise of the Impulse Response of a Random Linear Filter," I.R.E. Transactions on Information Theory, vol. IT-3, March, 1957, pp. 5-10.
14. Berman, M., Weiss, M. F., and Shahn, E., "Some Formal Approaches to the Analysis of Kinetic Delta in Terms of Linear Compartmental Systems," Biophysical Journal, vol. 2, 1960, pp. 289-316.
15. Berman, M., Shahn, E., and Weiss, M. F., "The Routine Fitting of Kinetic Data to Models," Biophysical Journal, vol. 2, 1962, pp. 275-287.
16. Bellman, R., Kagawada, H., and Kalaba, R., "Quasilinearization, System Identification and Prediction," The Rand Corporation, RM-3812-PR, Santa Monica, California, August, 1963.
17. Kumar, K.S.P., and Sridhar, R., "On the Identification of Linear Systems," 1964 Joint Automatic Control Conference, Stanford University, Stanford, California, June 24, 25, and 26, 1964, pp. 361-365.
18. Zadeh, L. A., and Desoer, C. A., Linear System Theory, McGraw-Hill Book Company, Inc., New York, New York, 1963.
19. Valkenburg, M. E. Van, Modern Network Synthesis, John Wiley and Sons, Inc., New York, N. Y., 1960, pp. 209-247.
20. Guillemin, E. A., Synthesis of Passive Networks, John Wiley and Sons, Inc., New York, N. Y., 1957, pp. 537-658.
21. Faddeeva, V. N., Computational Methods of Linear Algebra, Dover Publications, Inc., New York, N. Y., 1959.
22. Wilkinson, J. H., Rounding Errors in Algebraic Processes, Her Majesty's Stationary Office, London England, 1963.
23. Turing, A. M., "Round-off Errors in Matrix Processes," Quarterly Journal of Mechanics and Applied Mathematics, vol. 1, 1948, pp. 287-308.
24. Sarber, W. H. Jr., "On the Identification of Time-varying Systems," Identification Problems in Communication and Control Systems, Papers Delivered at the Fifty-fourth Meeting of the Princeton University Conference, March 21 and 22, 1963, pp. 177-201.
25. Booth, A. D., Numerical Methods, Butterworth, London, England, 1957, pp. 95-100, p. 158.
26. Householder, A. S., Principles of Numerical Analysis, McGraw-Hill Book Company, Inc., New York, N. Y., 1951.
27. Whittaker, E., and Robinson, G., The Calculus of Observations, London, Blackie and Son, Ltd., 1956, Chapter IX.

28. Laning, J. H., Jr., and Battin, R. H., Random Processes in Automatic Control, McGraw-Hill Book Company, Inc., New York, N. Y., 1956.
29. Rodman, R. D., A B-5000 ALGOL-60 Program for the Solution of an $N \times N$ Non-Homogeneous System of Linear Equations, Bulletin MRS-097, Burroughs Corporation, Detroit, Michigan, April 1, 1963.
30. Rodman, R. D., A Burroughs B-5000 ALGOL 60 Program for Finding all Roots of a General Function, Technical Bulletin MRS-122, Burroughs Corporation, Detroit, Michigan, September 15, 1963.
31. Pixley, A. F., and Macek, A. G., A B-5000 ALGOL 60 Runge-Kutta-Nystrom Starting Procedure for use With Milne Type Predictor-Correctors in Solving Systems of Ordinary Differential Equations, Technical Bulletin MRS-113, Burroughs Corporation, Detroit, Michigan, August 1, 1963.
32. Rodman, R. D., A B-5000 ALGOL 60 Program for the Evaluation of the Inverse of a Matrix, Technical Bulletin MRS-093, Burroughs Corporation, Detroit, Michigan, May, 1963.
33. Rodman, R. D., An Iterative Scheme for Reducing the Roundoff Error of the Inverse of a Matrix B-5000, Technical Bulletin MRS-096, Burroughs Corporation, Detroit, Michigan, May 1963.
34. Bechman, F. S., "The Solution of Linear Equations by the Conjugate Gradient Method," Mathematical Methods for Digital Computers, John Wiley and Sons, Inc., New York, N. Y., 1964, pp. 62-72.
35. Booth, A. D., Numerical Methods, Butterworth, London England, 1957, pp. 95-107.
36. Numerical Analysis, "Proceedings of Symposia in Applied Mathematics," vol. VI, McGraw-Hill Book Company, Inc., New York, N. Y., 1956.
37. Woodrow, P. J., Methods of Locating the Extreme of Multivariate Functions, Technical Report No. 4, Control Systems Laboratory, Princeton University, Princeton, N. J., Nov. 1962.
38. Langefors, B., "Ill-conditioned Matrices," Svenska Aeroplan Aktiebolaget, Linkoping Sweden, SAAB TN 22, June 10, 1953, pp. 11-13.
39. Johnson, C. L., Analogue Computer Techniques, McGraw-Hill Book Company, Inc., New York, New York, 1956.
40. Kunz, K. S., Numerical Methods, McGraw-Hill Book Company, Inc. New York, N. Y.

41. Powell, M. J. D., "An Iterative Method for Finding Stationary Values of a Function of Several Variables." Computer Journal, vol. 5, July 1962, pp. 147-151.
42. Wilde, D. J., Optimum-Finding Methods, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1962.
43. Crockett, J. B., and Chernoff, H., "Gradient Methods of Maximization," Pacific Journal of Mathematics, vol. 5, 1955, pp. 33-50.
44. McCracken, D. D., and Dorn, W. S., Numerical Methods and Fortran Programming, John Wiley and Sons, Inc., New York, N. Y., 1964.
45. Diamessis, J. E., "On the Determination of the Parameters of Certain Nonlinear Systems," (Correspondence). Proceedings of the Institute of Electrical and Electronic Engineers, vol. 55, March 1955, p. 319-320.
46. Diamessis, J. E., "A Method for Determining the Parameters of Certain Time Varying Systems," (Correspondence) Procedure of I.E.E.E., Vol. 53, April 1955, pp. 396-397.
47. Hildebrand, F. B., Introduction to Numerical Analysis, McGraw-Hill Book Company, Inc., New York, N. Y., 1956, pp. 436-439.
48. Dauldon, W. C., Variable Metric Method for Minimization, ANL-5990, Rev. 1, Argonne National Laboratory, Lemont, Illinois, Nov. 1959.

VITA

Clifford Overall Guffee Jr. was born in Madison, Tennessee, on December 10, 1936. He is the son of Clifford Overall Guffee, Sr. and Erlane Guffee.

He attended public school in Madison, and Goodlettsville, Tennessee, and graduated from Goodlettsville High School in 1955. He attended the University of Tennessee as a co-operative student with the U. S. Army Corps of Engineers, and graduated with a B.S.E.E., in June of 1960. From June of 1960 until September 1961 he held the position of Test Planning Engineer with Western Electric Company in Burlington, North Carolina. He entered graduate school at the Georgia Institute of Technology in September of 1961, where he held the position of Graduate Teaching Assistant until June of 1964. In June of 1964 he received his M.S.E.E. degree from the same school. He held a National Science Foundation fellowship during the summer of 1964, and a NASA fellowship from September of 1964 until September of 1965.